

2019

Biochemistry: Production of recombinant Rh blood group antigens for detection of alloimmunisation

Alkhanbashi, Mashael

<http://hdl.handle.net/10026.1/15239>

<http://dx.doi.org/10.24382/519>

University of Plymouth

All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Please cite only the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.



**UNIVERSITY OF
PLYMOUTH**

Biochemistry: Production of recombinant Rh blood
group antigens for detection of alloimmunisation

by

Mashaël Ahmed Alkhanbashi

A thesis submitted to University of Plymouth in partial fulfilment for the degree

of

DOCTOR OF PHILOSOPHY

School of Biomedical Sciences

July 2019

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with its author and that no quotation from the thesis and no information derived from it may be published without the author's prior consent.

Contents

List of Figures	7
List of Tables	11
Acknowledgements	12
Abstract	14
Author's Declaration	16
Abbreviations	17
Amino Acids	20
1 Introduction	21
1.1 Mammalian Red Blood Cells	22
1.1.1 Major RBC Membrane Proteins	23
1.1.1.1 Glycophorin	26
1.1.1.2 Band 3 Complex	27
1.1.1.3 Ankyrin	28
1.1.1.4 Spectrins	28
1.1.1.5 Protein 4.1 R	29
1.1.1.6 CD47	29
1.2 Blood group antigens	31
1.2.1 ABO System	35
1.2.2 RH System	36
1.2.3 Kell (KEL) System	36
1.2.4 MNS Antigen System	37
1.2.5 Gerbich System	38
1.2.6 Duffy System	38
1.2.7 Lutheran System	39

1.2.8	Kidd (JK) System	39
1.3	History of Rh Blood Group	41
1.4	Molecular Aspects of Rh	42
1.4.1	Rh _{null} Syndrome	46
1.5	Rh Proteins	48
1.6	<i>RH</i> polymorphisms	54
1.6.1	D-epitopes Models	57
1.7	D-Variant Phenotypes	60
1.8	Rh in Various Species	62
1.8.1	<i>Escherichia coli</i> (<i>E.coli</i>)	69
1.8.2	<i>Nitrosomonas europaea</i>	71
1.9	The Immunogenicity of Rh Blood Group	74
1.9.1	Haemolytic Transfusion Reactions	75
1.9.2	Haemolytic Disease of the Foetus and Newborn	75
1.10	Aims and Objectives	79
2	Materials and Methods	82
2.1	Materials	83
2.1.1	Primary Antibodies	83
2.1.2	Custom Primary Antibodies	83
2.1.3	Secondary Antibodies	86
2.1.4	Plasmids and Bacterial Strains	86
2.1.4.1	Plasmids and Expression vectors	86
2.1.4.2	Bacterial Strains	87
2.1.5	Erythrocyte samples	87
2.1.6	Producing Luria Bertani Media (LB Media)	87
2.1.7	Producing Luria Bertani Agar Plates (LB Agar)	88
2.2	Methods	89
2.2.1	Alignment of Rh for different species	89
2.2.2	Mapping of the Human RhD external loops and De- signing of the Recombinant Gene	89

2.2.3	Cloning	90
2.2.3.1	Cloning <i>NeRh50"6</i> Construct in pESV2 vector . .	90
2.2.3.2	Enzymatic Digestion	92
2.2.3.3	Purification of Digested DNA	93
2.2.3.4	Molecular cloning into pESV2	94
2.2.3.5	Transforming in <i>DH5α</i> <i>E. coli</i> competent cells . .	96
2.2.3.6	Isolation and purification of plasmid DNA	98
2.2.3.7	Confirmation tests	99
2.2.3.8	Cloning <i>NeRh50"6</i> Recombinant in pET expres- sion vector	107
2.2.3.9	Purification of Digested DNA	108
2.2.3.10	Molecular cloning into the pET Expression Vector	109
2.2.3.11	Transforming in <i>BL21(DE3)</i> <i>E.coli</i> competent cells	110
2.2.3.12	Isolation and purification of plasmid DNA	112
2.2.3.13	Confirmation tests	112
2.2.3.14	Protein Extraction	118
2.2.3.15	Sodium Dodecyl Sulphate-polyacrylamide Gel Electrophoresis (SDS-PAGE)	121
2.2.3.16	Flow cytometry	125
3	Results	128
3.1	Introduction	129
3.1.1	Aim of the study	132
3.2	Constructing <i>RH</i> recombinant genes for expression in <i>E.</i> <i>coli</i> membrane.	134
3.2.1	Mapping human RhD 6 th external loop	134
3.2.2	Purifying DNA	153
3.2.3	PCR amplifying the <i>NeRh50"6</i>	153
3.3	Expression Vectors	155
3.3.1	pESV2 expression vector	155

3.3.1.1	Digestion of pESV2- <i>NeRh50</i> "6	155
3.3.1.2	Cloning of pESV2- <i>NeRh50</i> "6^ Human <i>RHD</i> in <i>DH5α E. coli</i> Competent Cells	157
3.3.1.3	Sequencing of pESV2 expression vector	159
3.3.2	pET-a9 expression vector	166
3.3.2.1	Cloning of pET- <i>NeRh50</i> /Human RhD external loops in BL21 (DE3) Chemically Competent Cells	167
3.3.2.2	Sanger Sequencing of <i>NeRh50</i> "human <i>RhD</i> ex- ternal loops	169
3.4	Three-Dimensional Protein Model	176
3.5	Confirming the presence of the Rh Recombinant proteins . . .	180
3.5.1	Western Blot Results	181
3.5.1.1	Comparison of Bacterial Protein Extraction Methods	181
3.5.1.2	Detection of human RhD external loop 4 by 2.2B9 antibody by western blotting	183
3.5.1.3	Detection of human RhD external loop 6 by 2A4 / 5G9 antibodies by western blotting	185
3.5.1.4	Detection of human RhD external loops 4, 6 by BIRMA-D6 antibody by western blotting:	188
3.5.1.5	Detection of human RhD external loops 3, 4, 6 by MS-26 antibody by western blotting:	190
3.5.1.6	Detection of <i>NeRh50</i> external loop 1 and C- terminus by 1918 antibody by western blotting: . .	192
3.5.1.7	Detection of <i>NeRh50</i> external loop 1 by 1917 antibody by western blotting:	194
3.5.2	Flow cytometry Results	196
3.5.3	Transmission Electron Microscopy Results	205
3.6	Discussion	210
3.6.1	Generating constructs	210

3.6.2	Confirming the presence of constructs by testing for proteins expression	213
3.6.2.1	Western blot results	215
3.6.2.2	Flow cytometry results	219
3.6.2.3	TEM results	221
4	Discussion and Conclusion	223
4.1	General discussion	224
4.2	Expression of D-epitope in <i>E. coli</i>	229
4.3	Conclusions	232
4.4	Future Work	234
5	Appendices	261
5.1	Appendix A	262
5.2	Appendix B	268
5.3	Appendix C	274
5.4	Appendix D	281
5.5	Appendix E	288
5.6	Appendix F	295
5.7	Appendix G	302
5.8	Appendix H	309
5.9	Appendix I	316

List of Figures

Figure 1.1	Interaction between the Red Cell Membrane Components:	25
Figure 1.2	Red Cell Membrane Protein Components:	30
Figure 1.3	<i>RHD</i> and <i>RHCE</i> genomic structure in different status:	43
Figure 1.4	The predicted trimeric structure of the Rh complex.	50
Figure 1.5	Topology of the human RhD protein:	51
Figure 1.6	The 2-Dimensional structure of the human RhD and RhCE proteins.	53
Figure 1.7	Predictive models of D epitope configuration.	58
Figure 1.8	Secondary structure of RhCG protein.	65
Figure 1.9	Crystal structure of AmtB:	65
Figure 1.10	Sequence alignment of the Rh protein family members.	68
Figure 1.11	Topology structure of <i>E. coli</i> AmtB Channel:	70
Figure 1.12	The trimer structure of the Rh-like protein in <i>Nitrosomonas europaea</i> (<i>NeRh50</i>).	72
Figure 1.13	Topology structure of <i>NeRh50</i> protein in <i>N. europaea</i> :	73
Figure 2.1	Cloning and Transformation.	97
Figure 2.2	Screening process.	101
Figure 3.1	<i>Clustal X</i> alignment of the Transmembrane Domain (TM) (11-12) region of 13 Rh proteins from different species.	135
Figure 3.2	Amino acid sequence of Human RhD external loop 6 and the corresponding nucleotide sequence based on <i>E. coli</i> codon usage.	137

Figure 3.3	Predicted membrane topology and amino acid sequence of <i>NeRh50"6</i> construct.	139
Figure 3.4	pUC57 vector with the recombinant gene (<i>NeRh50"6</i>) circular sequence map.	141
Figure 3.5	pUC57 linear vector sequence with the recombinant gene (<i>NeRh5026</i>).	145
Figure 3.6	<i>BamHI</i> and <i>XbaI</i> digestion of pUC57- <i>NeRh50"6</i>	146
Figure 3.7	Amino acid sequence of Human RhD external loop 2 and the corresponding nucleotide sequence based on <i>E. coli</i> codon usage.	148
Figure 3.8	Amino acid sequence of Human RhD external loop 3 and the corresponding nucleotide sequence based on <i>E. coli</i> codon usage.	148
Figure 3.9	Amino acid sequence of Human RhD external loop 4 and the corresponding nucleotide sequence based on <i>E. coli</i> codon usage.	148
Figure 3.10	Predicted membrane topology and amino acid sequence of <i>NeRh50"4, 6</i> construct.	149
Figure 3.11	Predicted membrane topology and amino acid sequence of <i>NeRh50"3, 4, 6</i> construct.	150
Figure 3.12	Predicted membrane topology and amino acid sequence of <i>NeRh50"2, 3, 4, 6</i> construct.	151
Figure 3.13	pUC57 vector with the recombinant gene (<i>NeRh50"2, 3, 4, 6</i>) circular sequence map.	152
Figure 3.14	Amplicons of pUC57 plasmid - <i>NeRh50"6*</i>	154
Figure 3.15	<i>BamHI</i> / <i>NdeI</i> digestion of pESV2- <i>NeRh50"6</i>	156
Figure 3.16	PCR screening of positive pESV2- <i>NeRh50"6*</i> plasmid colonies, which amplified with AD1 / AD2 primers.	158
Figure 3.17	pESV2 vector constructed process.	159
Figure 3.18	The report of single sequencing run of the original (pESV2) vector sample library	163

Figure 3.19	The report of single sequencing run of the pESV2- <i>NeRh50"6</i> sample library	164
Figure 3.20	pET-9a expression vector map showing all restriction enzyme sites	166
Figure 3.21	Digested positive colonies of pET vector with the <i>NeRh50"6</i> , <i>NeRh50"4</i> , <i>6</i> , <i>NeRh50"3</i> , <i>4</i> , <i>6</i> , <i>NeRh50"2</i> , <i>3</i> , <i>4</i> , <i>6</i> constructs.	168
Figure 3.22	pET-NeRh50"2,3,4,6 vector map	171
Figure 3.23	Sanger sequencing data with electropherogram of all constructs cDNAs against (F1) forward primer.	172
Figure 3.24	Sanger sequencing data with electropherogram of all constructs cDNAs against (R1) reverse primer.	173
Figure 3.25	Sanger sequencing data with electropherogram of all constructs cDNAs against (F7) forward primer.	174
Figure 3.26	Sanger sequencing data with electropherogram of all constructs cDNAs against (R7) forward primer.	175
Figure 3.27	Predicted 3D protein model of <i>NeRh50" 2, 3, 4, 6</i> . . .	177
Figure 3.28	Predicted 3D protein model of the human RhD wild type	178
Figure 3.29	Predicted 3D protein model of <i>NeRh50</i> wild type . . .	179
Figure 3.30	Immunostaining of RBC membranes and <i>E. coli</i> pro- tein extractions to detect human RhD external loop 4 by using 2.2B9 antibody.	184
Figure 3.31	Immunostaining of RBC membranes and <i>E. coli</i> pro- tein extractions to detect human RhD external loop 6 by using 2A4 antibody.	186
Figure 3.32	Immunostaining of RBC membranes and <i>E. coli</i> pro- tein extractions to detect human RhD external loop 6 by using 5G9 antibody.	187

Figure 3.33	Immunostaining of RBC membranes and <i>E. coli</i> protein extractions to detect human RhD external loops 4, 6 by using BIRMA-D6 antibody.	189
Figure 3.34	Immunostaining of RBC membranes and <i>E. coli</i> protein extractions to detect human RhD external loops 3, 4, 6 by using MS-26 antibody.	191
Figure 3.35	Immunostaining of RBC membranes and <i>E. coli</i> protein extractions to detect NeRh50 external loop 1 and C-terminus by using 1918 antibody.	193
Figure 3.36	Immunostaining of RBC membranes and <i>E. coli</i> protein extraction using 1917.	195
Figure 3.37	MFI values of several Anti-D against various NeRh50/human RhD external loop constructs	200
Figure 3.38	MFI values of Anti- <i>NeRh50</i> (1918) against various NeRh50/human RhD external loop constructs	201
Figure 3.39	FITC fluorescence histogram of <i>NeRh50</i> human RhD external loops with Anti-D BIRMA-D10 (EpD8).	202
Figure 3.40	FITC fluorescence histogram of <i>NeRh50</i> human RhD external loops with Anti-D RUM-1 (EpD 6/7)	203
Figure 3.41	FITC fluorescence histogram of <i>NeRh50</i> human RhD external loops with Anti-D MAD-2 (EpD 19)	204
Figure 3.42	Transmission electron micrograph for a thin layer of <i>E. coli</i> contains recombinant <i>NeRh50</i> RhD by using anti-D loop 4 2.2B9.	207
Figure 3.43	Transmission electron micrograph for a thin layer of <i>E. coli</i> contains recombinant <i>NeRh50</i> RhD by using anti-D loop 6 2A4.	208
Figure 3.44	Transmission electron micrograph for a thin layer of <i>E. coli</i> contains recombinant <i>NeRh50</i> RhD by using anti- <i>NeRh50</i> loop 1 and C-terminus 1918.	209

List of Tables

1.1	Blood group systems as classified by the International Society of Blood Transfusion (ISBT).	34
1.2	The haplotype nomenclature of <i>RHD</i> and <i>RHCE</i> genes: . . .	55
2.1	Polyclonal(P)/ monoclonal(M) Anti-D and Anti- <i>NeRh50</i> antibodies	85
2.2	Primer sequences to add <i>NdeI</i> enzyme restriction site in the recombinant gene <i>NeRh50"6</i>	91
2.3	Reaction Setup for Barcoded Libraries	105
2.4	Primer sequences to amplify regions in pET- <i>NeRh50"2, 3, 4, 6</i>	114
3.1	List of human RhD external loops AA and <i>E. coli</i> nucleotide codons of these AA.	138
3.2	Anti-D and Anti- <i>NeRh50</i> antibodies list against different techniques.	180
3.3	Qubit protein quantitation results of three different colonies of pESV2- <i>NeRh50"6</i>	182
3.4	Flow Cytometric Analysis of wild type RhD, wild type <i>NeRh50</i> and 4 different constructs using Monoclonal Anti-D and Anti- <i>NeRh50</i>	198

Acknowledgements

First, I thank God for all blessings that enable me to complete this PhD project.

I would like to acknowledge and express sincere gratitude to the following people for their help during completion of my PhD; My director of studies, Dr. Tracey Madgett and my second supervisor Prof. Neil Avent and third supervisor Dr. Kris Jeremy for their devoted guidance, untiring patience, support- both inside and outside the lab, solicited and valued advice and their encouragement, without which this thesis would not be complete. We are grateful to Prof. Mike Merrick for providing (pESV2) plasmids.

A great thanks is also to Glenn Harper and Alexander Strachan Plymouth Electron Microscopy Centre (PEMC) where the electron microscopy preparation steps were done. I am also grateful to Paul Waines for his valuable help.

My husband and my colleague, Ibrahim Halawani, for his endless love, support and helpful discussions. It was impossible for me to finish this work without his support and encouragement. My family, my progeny Faisal and Wrd for their love, support, patience and understanding. Special thanks to my parents for their continued love, support and encouragement. I would like to thank all my friends and anyone who have been supporting me throughout my scholarship.

Finally, I would like to thank SACB for the fully scholarship,
without which this thesis would not be complete.



Abstract

Due to the importance of the antigen/ antibody reaction, all hospitals and blood banks worldwide rely on routine blood group phenotyping of samples. Using recombinant antigens removes the disadvantage of the current antibody screening when using the human red blood cells, which can complicate the detection of antibody in different cases: in the case of multi-transfused individuals and in the case of weak antigens. Although constructing recombinant antigens for some blood group antigens has been achieved successfully, such as; Kell, Duffy and Lutheran blood group antigens, no studies have been published describing successful production of *RH* recombinant antigens in a prokaryotic expression system. Rh proteins are conserved across many different species.

The aim was to construct a hybrid Rh protein comprised of a *Nitrosomonas europaea* (*N. europaea*) Rh50 backbone containing different external loops of the human RhD protein.

Human RhD external loop 6 was the most suitable candidate to start with due to its location between highly conserved domains of the protein. The hybrid gene for the *NeRH50*"human *RHD* was sub-cloned into an expression vector (pET) and transformed in BL21 (DE3) competent *E. coli*. We constructed three other hybrid genes with RhD human external loops on a *NeRh50* backbone: loops 4, 6; loops 3, 4, 6 and loops 2, 3, 4, 6.

Hybrid recombinant constructs were successfully expressed in the *E. coli* membrane as shown with western blot, flow cytometry, and transmission electron microscopy by the use of different commercial/ customized monoclonal and polyclonal antibodies against D epitopes in human RhD proteins and *NeRh50* like protein. Due to a lack of time, testing human sera against the recombinant antigens was not possible. However, the approach utilised in this study may lead to a new diagnostic assay whereby the expression of human RhD external loops in *E. coli* may be capable of detecting specific anti-Rh antibodies in patient serum.

AUTHOR'S DECLARATION

At no time during the registration for the degree of *Doctor of Philosophy* has the author been registered for any other University award without prior agreement of the Doctoral College Quality Sub-Committee.

Work submitted for this research degree at the University of Plymouth has not formed part of any other degree either at the University of Plymouth or at another establishment.

Name: Mashaël Ahmed Alkhanbashi

Signed:..... Date:.....

Word count of the thesis's main body: 40,248 words

Abbreviations List

A.A	amino acid
Amp	Ampicillin
BGG	blood group genotyping
bp	base pair
C	cytosine
cfDNA	cell-free DNA
DNA	deoxyribonucleic acid
dNTPs	deoxynucleotide triphosphates
EDTA	ethylenediaminetetraacetate
FHM	feto-maternal haemorrhage
FUC	Fucose
G	guanine
Gal	galactosyl
GalNAc	N-acetyl- D-galactosamine
gDNA	Genomic DNA
GDP-L-fucose	guanosine diphosphofucose
GTA	3--N-acetylgalactosaminyltransferase
GTB	3--galactosyltransferase
HDFN	haemolytic disease of the foetus and newborn

HEA	human erythrocyte antigen
HTR	haemolytic transfusion reaction
IAT	indirect antiglobulin test
IgG	immunoglobulin G
IgM	immunoglobulin M
ISBT	International Society of Blood Transfusion
ISPs	ion sphere particles
kb	kilobases
LR-PCR	long-range polymerase chain reaction
mRNA	messenger RNA
MFI	mean fluorescence intensity
NCBI	National Centre for Biotechnology Information
NGS	next-generation sequencing
PCR	polymerase chain reaction
RBC(s)	red blood cell(s)
RFLP-PCR	restriction fragment length polymorphism- polymerase chain reaction
RNA	ribonucleic acid
RT-PCR	real time PCR
SBT	sequencing-based typing
SCD	sickle cell disease
SNPs	single nucleotide polymorphisms

SNV single nucleotide variant

TEM Transmission Electron Microscopy

TM Transmembrane

Amino Acids List

Alanine	Ala
Arginine	Arg
Asparagine	Asn
Aspartic acid	Asp
Cysteine	Cys
Glutamic acid	Glu
Glutamine	Gln
Glycine	Gly
Histidine	His
Isoleucine	Ile
Leucine	Leu
Lysine	Lys
Methionine	Met
Phenylalanine	Phe
Proline	Pro
Serine	Ser
Threonine	Thr
Tryptophan	Trp
Tyrosine	Tyr
Valine	Val

Chapter 1

Introduction

1.1 Mammalian Red Blood Cells

Red blood cells (RBCs) are bright red, doughnut shaped, enucleated eukaryotic cells present in the peripheral circulation and circulate the body to supply oxygen (O_2) to the organs and transport carbon dioxide (CO_2) to the lungs, and are the most numerous cell type in the body. Moreover, RBCs are able to transport nutrients, waste products, and hormones. RBCs are produced in the Bone Marrow (BM) and stay in the circulation for about 120 days (Yawata, 2003).

The plasma membrane or cell membrane of RBCs (erythrocytes), which is a thin semi-permeable membrane is easy to isolate and use as a model to study the membrane structure and to perform biochemical analyses for the membrane's component, due to the lack of both nuclei and internal membranes (Gorter & Grendel., 1925).

The red cell membrane is highly dynamic and has a complicated structure. It is composed of an asymmetrical lipid bilayer, which makes 40% of the red cell membrane, almost 10% carbohydrates, and 50% are proteins both transmembrane proteins, and a filamentous meshwork of proteins, which forms a membrane skeleton along the entire cytoplasmic surface of the plasma membrane.

The 40% lipids of the red cell membrane are classified into phospholipids (PL) (60%), neutral lipids (30%) mainly cholesterol,

and glycolipids (10%). Phosphatidylcholine (PC) (31%) is the major phospholipid in the red cell membrane. However, phosphatidylethanolamine (PE) (29%), sphingomyelin (SM) (26%) and Phosphatidylserine (PS) (13%) are presented in the red cell membrane (Yawata, 2003).

The phospholipids play a role in the structure of the membrane and are thought to be important in the maintenance of red cell shape. The most important thing about phospholipids is that PS is located in the inner leaflet of the membrane in normal haemostasis, but it will switch to the outer leaflet of the membrane in non-healthy conditions such as aged, infected, or sensitised red cells (Pinder *et al.*, 1978; Lang *et al.*, 2012)

1.1.1 Major RBC Membrane Proteins

Due to the presence of the long twisted alpha and beta spectrin strands and actin filaments, the plasma membrane is linked to the RBC's cytoskeleton (Pinder *et al.*, 1978). Moreover, another link is provided by band 4.1, which binds to the junctions of spectrin and actin, as well as to glycophorin proteins.

The importance of the RBCs component interaction is to achieve the critical functions of the RBCs such as; transporting electrolytes from the RBCs to the plasma or vice-versa, and maintaining the RBCs shape and their elasticity and stability. There are two ways to show the interaction between the important

RBCs proteins component on the cell membrane: vertical interaction, which mediated by protein 4.1R that couples the junctional complex to Glycophorin C and horizontal interaction, which is mediated by ankyrin that couple spectrin to Band 3 (Figure 1.1).

Figure has been removed due to Copyright restrictions.

Figure 1.1. Interaction between the Red Cell Membrane Components:

This figure shows how these proteins interact with each other to achieve the critical functions of the RBCs such as; transporting electrolytes from the RBCs to the plasma or vice-versa, and maintaining the RBCs shape and their elasticity and stability. There are two types of interaction within the Red cell membrane; a vertical interaction is the main reason for the extreme elasticity of the red cell membrane. This type of a interaction occurs between Band 3, and the Rh core complex has interactions with Ankyrin. The other kind is the horizontal interactions where the spectrin dimers and protein 4.1R with phosphatidylserine, a cytoplasmically localised phospholipid, also occur.

Adapted from Avent (2018).

To demonstrate the importance of the transmembrane proteins and the interactions within the plasma membrane and the cytoskeleton components, any defect or mutation in them will affect both the shape and the function of the RBCs and can lead to many disorders, such as anaemia.

RBCs membrane proteins are classified into three different categories based on their function. Firstly, integral proteins,

these proteins are firmly attached to the membrane and act as the main transport channel such as; Glycophorin and Band 3 protein.

Secondly, peripheral proteins, cortical cytoskeleton components, which lie underneath the membrane and form a meshwork all over the membrane and determines cell shape, the most abundant; spectrin. Other peripheral membrane proteins of red blood cells; actin, ankyrin, and band 4.1R.

Thirdly, the anchoring proteins, which maintain the membrane structure such as Ankyrin (Von Heijne, 2006) (Figure 1.2).

1.1.1.1 Glycophorin

Glycophorin is a small integral glycoprotein, consisting of 131 amino acids with a molecular weight of about 30,000. Glycophorins were one of the first identified transmembrane proteins, they play important roles in maintaining and stabilizing the membrane shape of RBCs. Moreover, glycophorin is rich of sialic acid which gives the surface of RBCs a polar charge that assists in the movement of RBCs.

Glycophorins A and B are single transmembrane α helix of 23 amino acids, it is heavily glycosylated, with oligosaccharides, which exposed on the plasma cell surface, attached to 16 sites on the extracellular portion of the polypeptide chain.

Glycophorin C (GPC) and glycophorin D (GPD) interact with

protein 4.1R, contributing stability to RBC membrane (Alloisio *et al*, 1993; Marfatia *et al*, 1994).

1.1.1.2 Band 3 Complex

Band 3 is a glycoprotein and considered the major integral protein in the membrane and it is known as anion exchanger (AE1), with a molecular mass of 100 kDa on SDS-PAGE. Band-3 is composed of 911 amino acids and the gene (*EPB3*), which is responsible for AE1, is located on chromosome 17.

There are 1-2 million copies of AE1 per red blood cell membrane, it consists of three domains, an N-terminal domain, a short C-terminal domain, and membrane-spanning domain acts as a chloride/ bicarbonate exchange anion exchange channel (Bruce *et al.*, 2003).

Band 3 carries out the one to one exchange of bicarbonate (HCO_3^-) for Chloride (Cl^-) ions, transport of which can be reversed due to pH changes (Bohr shift). HCO_3^- is transported into the red cell in metabolising tissues, and out of the red cell in the lungs.

However, the band 3 complex not only attaches to the lipid bilayer membrane, but also interacts with other major membrane protein such as Ankyrin, Spectrins, and Protein 4.1R to maintain the structure of RBCs.

1.1.1.3 Ankyrin

Ankyrin is one of the primary anchoring membrane proteins in the RBC membrane. There are 1.2 million copies in RBC. It has a molecular weight of 206 kDa on SDS-PAGE. Ankyrin is a polar protein and is involved in the local segregation of integral membrane proteins. It contains three domains; N-terminal domain that is connected to AE1 while the heavily phosphorylated central domain is connected to spectrins and an acidic C-terminal domain (Bennett & Baines., 2001).

Ankyrin plays an important role in linking the plasma membrane with the cytoskeleton by binding to both spectrin and band 3.

1.1.1.4 Spectrins

Spectrins are the largest proteins of the RBCs membrane skeleton, composing of 75% of the membrane skeletal proteins and present at a concentration of about 200,000 copies per red cell. Spectrins are composed of two subunits, α -Spectrin and β -Spectrin. Each one of the spectrins is located on a different chromosome and has a different gene. α -Spectrin gene is *SPTA1* and located on chromosome 1 at position 1q22-q23 while β -Spectrin gene is *SPTA 2* and located on chromosome 14 at position 14q23-q24.2.

Spectrin proteins form the inner shell of the RBC membrane and are considered to be one of the major determinants of RBC

shape due to their flexibility, which can change the RBC membrane shape and so, any defect in this protein, leads to abnormal RBC membrane shape and some blood disorders.

1.1.1.5 Protein 4.1 R

Protein 4.1R is composed of 588 amino acids and is present at the concentration of approx. 200,000 copies per red cell. Protein 4.1R is an essential phosphoprotein within cell membranes. It has a molecular weight of 66 kDa, but on SDS-PAGE it migrates to 80 kDa. Protein 4.1R gene (*EPB41*) is located on chromosome 1p36.1 (Delaunay 1995; Yawata, 2003).

1.1.1.6 CD47

CD47 is also known as Integrin-associated protein (IAP). CD47 is a 47-52 kDa transmembrane receptor with five membrane spanning segments. It is a member of the immunoglobulin (Ig) superfamily and has a highly glycosylated extracellular immunoglobulin variable domain (IgV) at the N-terminus and five isoforms at the C-terminus. CD47 acts as a marker of self on the surface of RBCs to prevent phagocytosis by binding to Signal Regulatory Protein α (SIRP α). On senescent RBCs, CD47 levels drop and the immune system engulfs these cells and clears them from the circulation (Avice *et al.*, 2001).

Figure has been removed due to Copyright restrictions.

Figure 1.2. Red Cell Membrane Protein Components:

This figure shows part of the important Red Cell membrane components and membrane skeleton. The major RBC membrane components are Band 3 and Glycophorin C (GPC) they appear to be in higher quantity (1,100,000 and 200,000 copies per red cell) respectively, than the other protein members. Next, the Rh trimer complex, which consists of two RhAG and one RhD or RhCE, this complex has approximately 100,000 copies per red cell.

Adapted from eClinPath.com, Cornell University,
<http://eclinpath.com/hematology/physiology/erythrocytes/f2-large/>

1.2 Blood group antigens

In 1901, blood groups were identified by an Austrian scientist, called Karl Landsteiner. Moreover, he developed a blood group classification system based on the presence of agglutinin factors in the blood. This test was based on his observation of clump formation in the individual isolated plasma, which identified later as ‘agglutination’ in a paper titled as Agglutination phenomena of normal human blood, which led to identify three antigens of the ABO blood group system (A, B, and O). Later in 1902, a fourth antigen (AB) was discovered by Decastello and Sturli, respectively (Landsteiner, 1961, Levine, 1961, Schwarz and Dorner, 2003, Giangrande, 2000). Furthermore, in 1937, the blood transfusion patient’s death rate were reduced due to the discovery of Rhesus factor by Landsteiner and the help of another scientist, Alexander S. Wiener (Landsteiner & Wiener, 1940; Landsteiner & Wiener, 1941).

The blood group systems and their antigens are classified according to *The International Society of Blood Transfusion (ISBT)* nomenclature system. Currently, there are 39 blood group systems (Table 1.1).

Erythrocyte membrane components are critical in defining the blood group systems, each of the blood groups are characterised by specific antigens, which can be defined as inherited polymorphic markers present on the RBCs external membrane and can

be detected serologically by specific antibodies. These antigens can trigger an individual's immune response by the synthesis of corresponding antibodies, which occur normally after the individual is exposed to antigens; either from the environment or alloimmunisation during blood transfusion or pregnancy for individuals who lack these antigens (Reid et al., 2012, Daniels, 2013).

These antigens are either protein or carbohydrate molecules attached to protein or lipid of the RBC membrane forming protein, glycoprotein, or glycolipid structures (Reid and Mohandas, 2004; Daniels, 2005). There are two major classifications in order to define blood groups according to their antigens. Firstly, the carbohydrate antigens, which are synthesised by glycosyltransferases enzymes that are responsible for transferring a monosaccharide moiety from an active nucleotide sugar, which is called a glycosyl donor to nucleophilic glycosyl acceptor molecule. Secondly, the protein antigens, where any alteration or changing the sequence or configuration of the protein or indeed the complete loss of RhD and Fy, which is the Duffy antigen, is critical and causes recognition as an antigenic determinant.

Five blood group systems are defined by carbohydrate structures (ABO, H, P1Pk, I, GLOB); two are obtained from the plasma (LE, CH RG). The remaining are characterised by the protein sequence of the RBC membrane proteins. Moreover, major pro-

teins are: (DI, RH, RHAG, MNS, GE, and CO) most of them are expressed, and their function is membrane transporters, except MNS and GE which function as sialic acid carriers and membrane receptors (Table 1.1).

In the following section some of the major and minor blood group antigens are discussed, given their importance in transfusion and haemolytic disease.

Figure has been removed due to Copyright restrictions.

Table 1.1. Blood group systems as classified by the International Society of Blood Transfusion (ISBT).

1.2.1 ABO System

The first blood group system discovered and the most important blood group in transfusion and transplantation. This system is based on the carbohydrate antigens (N-Acetylgalactosamine and galactose).

This system can be divided into four categories: A, B, AB, and O. The individuals with B group have antibodies against A present in their blood circulation, and individual's with A group have antibodies against B present in their blood circulation. While people with group AB have both A and B antigens but no antibodies in their serum. Moreover O individuals have H antigens and have both antibodies in their blood circulation. These antigens can be expressed on the RBCs membrane additionally to some other cells (Anthea *et al.*, 1993, Yamamoto, 2004, Storry and Olsson, 2009).

Anti-A and anti-B are produced in the early years of life. The first explanation of presence for these antibodies, is as a consequence of exposure to environmental substances, such as food, bacteria, and viruses and the production of these antibodies is due to the similarity of these environmental substance epitopes to A and B glycoprotein antigens. Moreover, the other possible explanation of the presence for anti-A and anti-B in individual's serum is due to the presence of A and B antigens in human gut flora and, thus leading to the production of anti-A and anti-B (Carel, 2004).

Early 1980s, at the New York Blood Center, Goldstein and his colleagues successfully created a H group antigen from blood group A and B by using a specific exo-glycosidases capable of removing the immunodominant sugar residues. Thus, able them to reduce the mismatching error that happened after having an incompatible blood transfusion, which lead to fatal transfusion reactions, moreover, creating a universal blood supply.

In 2000, Kruskall and his colleagues were successfully utilizing a recombinant enzyme-converted group-B-to-group-O (ECO). They conclude that ECO RBCs were comparable to group O cells for safety and efficacy.

1.2.2 RH System

The RH system is the second most important blood group system after ABO system. It is based on protein antigens; most major antigen is D. It can cause Haemolytic disease of the fetus and newborn (HDFN) when the blood of the mother is incompatible with her fetus, and haemolytic transfusion reactions (HTR). In this thesis we will discuss this blood group system in more depth. See later in introduction.

1.2.3 Kell (KEL) System

In 1946, Kell blood group system was named after Mr. Kelleher, who was the first producer of its antibody (Coombs *et al.*, 1946). Kell blood group system is considered to be the third most clin-

ically important blood group system and the sixth blood group system (006) according to the ISBT. The Kell glycoprotein is a transmembrane, single-pass protein with an enzymatic function that carries the Kell antigens. This system is highly polymorphic, it consists of 35 antigens, which all expressed on the surface of the Kell glycoprotein on the RBCs membrane, on haematopoietic tissue and some other tissues such as; brain, lymphoid organs, heart and skeletal muscle (Kaita *et al.*, 1959; Russo *et al.*, 2000; Daniels *et al.*, 2007; Reid *et al.*, 2012).

Anti-K (KEL¹) causes severe HDFN and HTR (Vaughan *et al.*, 1998; Daniels *et al.*, 2003). Additionally to other antigens including; k (KEL²), Kpa (KEL³), Kpb (KEL⁴), Jsa (KEL⁶) and Jsb (KEL⁷), which have the ability to stimulate the production of haemolytic antibodies and causing HDFN (Daniels, 2002).

1.2.4 MNS Antigen System

In 1927, Landsteiner and Levine described the MNS Antigen system, which consists of 48 antigens. The responsible genes are: Glycophorin A and Glycophorin B. Anti-M and anti-N antibodies are usually IgM types and rarely associated with transfusion reactions.

1.2.5 Gerbich System

Glycophorin C (GPC), 128 amino acid residues and glycophorin D (GPD), 107 amino acid residues, which are both encoded by a single gene. *GYPC*, is the responsible protein to express the Gerbich blood group system with 11 antigens. The high dominance antigens are; Ge2, Ge3, Ge4, GEPL [Ge10*], GEAT [Ge11*], GETI [Ge12*] (Alloisio *et al.*, 1993; Reid & Spring, 1994; Marfati *et al.*, 1994).

Gerbich antigens act as receptors for the malarial parasite *Plasmodium falciparum*. Anti-Ge2 and anti-Ge3 have caused hemolytic transfusion reactions, and anti-Ge3 has produced HDFN.

1.2.6 Duffy System

The Duffy System is also known as Fy glycoprotein, CD234. It presents and is expressed on the surface of RBCs and other tissues such as; kidney, lung, liver, spleen, brain (Iwamoto *et al.*, 1996) and colon (Chaudhuri *et al.*, 1997). Antigens Fy^a and Fy^b, which are located on the Duffy glycoprotein, can result in four possible phenotypes, namely Fy(a⁺b⁻), a Fy(a⁺b⁺), Fy(a⁻b⁺), and Fy(a⁻b⁻). The antibodies against this system are predominantly IgG subtype (Tournamille *et al.*, 1995).

The Duffy glycoprotein is a receptor for cytokines, which are chemicals that are secreted by white blood cells during inflam-

mation. Moreover, it acts as a receptor for *Plasmodium vivax*, a parasite causes malaria. RBCs with Duffy antigens deficiency are relatively resistant to invasion by *Plasmodium vivax*.

Duffy blood group system antibodies can cause haemolytic transfusion reactions (Boyland *et al.*, 1982; Sosler *et al.*, 1989) and HDFN (Vescio *et al.*, 1987; Goodrick *et al.*, 1997).

Ridgwell *et al.*, in 2007 successfully produced soluble recombinant blood group antigens; Duffy blood group antigen was one of them. In this study they prove that the soluble recombinant blood group antigens can mimic blood group antigens and react with human allogeneic antibodies.

1.2.7 Lutheran System

Lutheran (Lu, B-CAM, CD239) blood group antigens are carried by membrane glycoproteins (Campbell *et al.*, 1994), which are a member of the immunoglobulin gene superfamily (IgSF) (Parsons *et al.*, 1995). It is expressed in RBCs membrane and other tissues (Reid & Lomas-Francis, 2012).

Lutheran blood group antigens are involved in a mild type of haemolytic transfusion reactions (Daniels, 2002) and rarely involved in HDFN (Inderbitzen *et al.*, 1982).

1.2.8 Kidd (JK) System

Kidd antigens, which are also known as Jk antigen is a glyco-

protein. The Kidd protein is expressed on the RBCs membrane and kidney tissues and acts as a urea transporter in RBCs and renal endothelial cells. This system consists of 3 antigens: Jk_a, Jk_b and Jk3.

Although Kidd antibodies are rare, they can cause severe transfusion reactions.

1.3 History of Rh Blood Group

In 1939, the Rh system was clinically described in a historic paper written by Levine and Stetson for a case of miscarriage, where the fetus had been lost due to a haemolytic reaction after a blood transfusion from her husband. They assumed that the mother's serum was immunised by her fetus, which had inherited this newly described antigen from their father.

In 1940, Rhesus red cell factor was described by Landsteiner and Wiener (Landsteiner & Wiener, 1940; Landsteiner & Wiener, 1941). It is considered as one of the most immunogenic and polymorphic blood grouping systems (Levine *et al.*, 1941). In fact, Landsteiner & Wiener had described anti-LW which shows different degrees of reactivity against RhD positive and negative red cells.

The Rh system comprises 56 antigens, five of them are the most important; (D, E, e, C, c) (Conroy *et al.*, 2005). The presence or absence of the D antigen from RBCs surface can distinguish between Rh positive (in the presence of D antigen) or Rh negative (in the absence of D antigen) (Colin *et al.*, 1991). Moreover, the *RHCE* gene is related to C/c and E/e epitopes, and they represent other polymorphisms of the Rh complex.

1.4 Molecular Aspects of Rh

The Rh family includes three erythroid gene members: *RHD*, *RHCE*, and *RHAG* (Matassi *et al.*, 1999) and two non-erythroid proteins RhBG and RhCG (Huang *et al.*, 2000). The erythroid genes are involved in forming the essential complex of Rh antigen expression and red blood cell membrane integrity (Huang, 1998). *RHD* and *RHCE* are located at the *RH* locus on chromosome 1p34-1p36 (Cherif-Zahar *et al.*, 1991; Matassi *et al.*, 1999), while *RHAG* is present on chromosome 6p12- 6p21 (Cherif-Zahar *et al.*, 1991; Matassi *et al.*, 1999).

Three genes are expressed in the erythroid lineages (Van Kim *et al.*, 2006). *RHD* positive locus is composed of not only *RHD* but also the *RHCE* gene (Van Kim *et al.*, 2006). Both *RHD* and *RHCE* genes have ten exons, share comparable exon/intron organisation, and can encode different proteins (Matassi *et al.*, 1999; Westhoff, 2004). genes,

The differences between *RHD* and *RHCE*, although they are located on the same chromosome, they are orientated in opposite directions (*RHD* 5' to 3', *RHCE* 3' to 5'). They carry different antigens (D, CE), respectively (Van Kim *et al.*, 2006) (Figure 1.3).

Moreover, Rh family include other two non-erythroid members, which are *RHBG* and *RHCG* which are mainly expressed in the liver and kidney, one of their functions is related to ammonia

genesis and excretion (Avent *et al.*, 2006; Van Kim *et al.*, 2006).

Figure has been removed due to Copyright restrictions.

Figure 1.3. *RHD* and *RHCE* genomic structure in different status:

RHD and *RHCE* genes are located on chromosome 1p36.11, with 10 exons of each and opposed direction *RHD* 5' to 3' while *RHCE* is 3' to 5'. A it shows both *RHD* and *RHCE* genes in the D positive individuals. In B the *RHD* pseudo-gene (*RHD* ψ) it shows a 37bp duplication insert between exon 3 and 4 of *RHD*. There are three missense mutations in exon 5 and a nonsense mutation in exon 6, which stop the protein being expressed. In this case, the individuals still have normal *RHCE* gene. In C, D-negative shown where there is a complete deletion of the *RHD* gene with normal *RHCE* gene. This case commonly appears in Caucasian.

Adapted from Avent, 2018

In 2000, Singleton *et al.* studied 98 blood samples from a mixed race of black African donors to screen for the *RHD* pseudogene (*RHD ψ*) (Figure 1.3). Three samples out of the 98 samples were D-negative samples while the rest were D-positive samples. The screening results show 14 samples out of the total number of samples have the *RHD ψ* . Finding the difference between the *RHD* and *RHD ψ* , the exons and introns of both genes were sequenced.

The result of this sequencing of some D-negative Africans samples showed the existence of the *RHD* gene, but no mRNA was produced due to an additional 37bp insertion, which causes a stop codon at position 210. This 37 bp is a duplication of a sequence spanning the intron3- exon 4 boundary.

Secondly, the sequencing of the *RHD ψ* exons 1, 3-7, and 10 showed more alterations within the *RHD* gene. For example exon 4 has a mutation from point G to A at nucleotide 609.

Thirdly, the sequencing showed some missense mutations in exons 3 and 5 encoding Met 218 Ile, Phe 223 Val, and Ser 225 Phe substitutions. Fourthly, in exon 6, the sequencing showed an alteration in the nucleotide 807 from T to G, which translates as a stop codon.

Finally, Singleton *et al.* concluded that the inactive *RHD* gene would be referred to here as *RHD ψ* . Moreover, there are more conclusions in this study such as; they found that about two-thirds of D-negative Africans have an inactive *RHD* gene.

While the remaining third is divided into two groups; one characterised by homozygous *RHD* deletion and the other group is characterised by having the *RHD-CE-Ds* hybrid gene, which is capable of producing c, C, e but not D. Also, the D-negative African Americans and South African people of mixed race are having the same 3 genetic backgrounds, but with different percentage.

24% of African Americans have the *RHD* ψ , and 17% of South African donors of mixed race have *RHD* ψ , while 81% of South African donors of mixed race and 54% of African Americans, 81% of South African donors of mixed race have no *RHD*. Singleton *et al.*, 2000 found that 66 percent of donors had the *RHD* ψ , 18% completely lacked the *RHD*, and 15 percent had the hybrid *RHD-CE* gene that is linked to the positive and negative phenotype.

Moreover, most *RHD* negative cases in Caucasians (Mourant *et al.*, 1976; Tills *et al.*, 1983) are caused by a deletion of the *RHD* gene with no molecular change in the *RHCE* gene (Colin *et al.*, 1991; Faas *et al.*, 1995; Avent *et al.*, 1997).

In non-Caucasians, the individuals can have the *RHD* negative phenotype while they carry the entire *RHD* gene (Daniels *et al.*, 1997; Okuda *et al.*, 1997).

1.4.1 Rh_{null} Syndrome

Rh_{null} is an inherited haemolytic disease, where the RBCs look abnormal morphologically and are dysfunctional (Schmidt *et al.*, 1967; Schmidtl *et al.*, 1969). Rh_{null} can cause some degree of haemolytic anaemia due to the abnormality in the organisation of their RBCs membrane phospholipids (Kuypers *et al.*, 1984). Rh_{null} RBCs membrane shows more cation permeability than the normal RBCs (Lauf *et al.*, 1976).

Rh_{null} has two types; Amorph type and Regulator type; and it's related to completely deficient in RhD and RhCE proteins. The deficiency caused by molecular changes (silencing mutations) in *RHCE* associated with deletion of *RHD*, which lead to the lack of Rh proteins expression (Chown *et al.*, 1972). However, *RhAG* is still expressed in small amounts (Cherif-Zahar *et al.*, 1998; Huang *et al.*, 1998). Regulator type is related to a molecular defect in *RHAG* gene (Chérif-Zahar *et al.*, 1996).

The importance of the *RHAG* in the regulator type is the association of the presence of this gene with the expression of Rh antigens (Chérif-Zahar *et al.*, 1996; Avent *et al.*, 2000).

There is an interaction between the Rh protein complex and other protein members of the RBC membrane such as Glycophorin B (GPB), CD47 or integrin-associated protein (IAP), and Landsteiner-Wiener glycoprotein (ICAM-4)(LW) (Cartron *et al.*, 1994).

Both Rh_{null} types have shown a reduction in GPB (Dahr *et al.*,

1987; Avent *et al.*, 1988), is considered as a major sialoglycoproteins RBCs membrane component can be linked to the MNS and SC blood groups respectively.

According to Von dem Borne *et al* (1990), they noticed the different in the Rh glycoprotein N-glycan chains size between the the normal RBCs and the Rh_{null} RBCs by using anti-RhAG MoAb 2D10 with co-precipitation techniques. In normal RBCs that were GPB-deficient, the Rh glycoprotein N-glycan chains size were increased while the size of this chain was reduced in the Rh_{null} RBCs. The molecular weight difference is due to faster moving of the Rh glycoprotein chain in the Rh_{null} RBCs through intracellular membranes to the cell surface. Moreover, they confirmed that both GPB and the Rh30 polypeptides interact with the Rh glycoprotein during the biosynthesis of the Rh complex.

The RBC membrane of the Rh_{null} individual has shown a the reduction of CD47 compared to healthy RBC membranes (Miller *et al.*, 1987; Avent *et al.*, 1988; Lindberg *et al.*, 1994). The role of CD47 is associated with the prevention of phagocytosis of RBCs. It functions as a self-marker by binding to signal regulatory protein an alpha (SIRP a) on macrophages, to create a negative signal that can reduce the “eat me” signal and protect the RBCs from engulfment (Oldenborg *et al.*, 2000).

According to some studies, in Rh_{null} patients, NH₃ permeation cannot be measured directly from the RhD and RhCE due to the

absence of the RhAG; both RhD and RhCE cannot be expressed. Furthermore, there are two different experiments show the role of the RhAG in transporting the NH_4^+ (Westhoff, 2006).

1.5 Rh Proteins

The Rh protein family is divided into two main categories; non-glycosylated and glycosylated (Eyers *et al.*, 1994).

Non-glycosylated protein group contains; RhD and RhCE with 417 amino acids (416 amino acid mature protein as the N-terminal methionine is cleaved), which associate with another protein located in the red blood cell membrane such as (RhAG) (Conroy *et al.*, 2005).

The molecular weight of RhD and RhCE proteins is 30 kDa, while the (Rh-50) consist of RhAG glycoprotein which measured 40-100 kDa on SDS-PAGE with 409 amino acid residues (Eyers *et al.*, 1994). RhCe and RhD are very homologous; they differ in just 35-36 amino acids (Avent *et al.*, 1990; Cherif-Zahar *et al.*, 1990).

The important region in the RhD protein, which has the conformational difference to the RhCE protein is, the predicted third, fourth and sixth external loops of the RhD polypeptide (Liu *et al.*, 1999 a). The alteration on these amino acids will affect the creation of the D epitopes.

The Rh complex (RhD, RhCE, RhAG, and CD47) is located

embedded in the phospholipid bilayer of the RBCs membrane. Rh_{null} cases were studied in order knowing the location and the function of this complex (Cartron, 1999; Huang *et al.*, 2000).

Moreover, it shows that the Rh complex is connected to the actin-spectrin-based RBCs membrane skeleton, presumably via RhD, RhCcEe, RhAG and/or CD47 (Figure 1.1).

In 1994, Eysers's group proposed that erythrocyte Rh assembly is a RhAG2 Rh-30 2 tetramer. The tetrameric model of Rh complex is created by the interacting of RBCs membrane glycoproteins components such as LW, CD47 and GPB (Cartron, 1994).

In 2005, Conroy *et al.*, proposed a model of the human Rh proteins, RhD and RhAG by using one of the *Escherichia coli* ammonia channel AmtB proteins (ammonia transporters) in bacteria as a template. Moreover, they used the predicted tertiary structure of Rh proteins and the available biochemical data for the Rh proteins.

The conclusion of this study leads to suggest that the erythrocyte Rh complex is trimeric (Figure 1.4).

According to the model, they suggested that RhAG and the homologous non-erythrocyte Rh glycoproteins, RhBG and RhCG, have very similar channel architecture to AmtB.

According to Rh topology prediction, it is believed the Rh proteins have 12 transmembrane spanning domains with intracellular N and C-termini, unlike the AmtB member family, which has

Figure has been removed due to Copyright restrictions.

Figure 1.4. The predicted trimeric structure of the Rh complex.

This structure was predicted based on the *E. coli* AmtB and human RhCG crystal structures. It is composed of 2 molecules of RhAG and one molecule either RhD or RhCcEe. The dimension of each monomer is around 46 Å. The red residues show the critical amino acid in the epitope of external loop 4, while the purple residues show them in external loop 6. The transmembrane α -helices are numbered from (M0 to M11).

Adapted from Callebaut *et al.*, (2006).

11 transmembrane domains (Cherif-Zahar *et al.*, 1990; Avent *et al.*, 1992) (Figure 1.5).

Figure has been removed due to Copyright restrictions.

Figure 1.5. Topology of the human RhD protein:

Twelve transmembrane domains are shown with six external loops, the arrows show the 2, 4, 6 (left to right) external loops of the human RhD and their critical amino acid. The black coloured circle shows the equivalent residues to the His those present in the AmtB (Tyr173, Phe316). The red circle at position (353Gly) is representing the proteolytically susceptible area. The cyan coloured circles are related to the critical position residues in the channel. The yellow coloured circle is for the surface exposed (Cys). The orange, purple, green, and blue coloured circles show the epitopic residues. Adapted from Conroy *et al.*,(2005).

In RBCs, the most important probable function of RhAG is to detoxify the organs: such as the liver and the kidney by transporting NH_3 and/or NH_4^+ .

The role of RhAG in the non-erythroid tissues would be to regulate the acid-base balance (Van Kim *et al.*, 2006). Comparing the Rh proteins models: the RhAG protein model has a central region occupied by histidine amino acids. Although, in both RhD and RhCE proteins, these histidines are replaced by other amino acids (tyrosine and phenylalanine); histidine number 185 is substituted with tyrosine and histidine number 344 is replaced

by phenylalanine (Javelle *et al.* 2006).(Figure 1.6).

Figure has been removed due to Copyright restrictions.

Figure 1.6. The 2-Dimensional structure of the human RhD and RhCE proteins.

The blood group active Rh proteins, RhD and RhCcEe both have 12 membrane spanning segments. It shows the similarity between them and the difference. Both of them comprise of 10 exons, and they share similar amino acid except for (35 or 36 depending on RhCE status) changes between RhD and RhCcEe. The yellow circles in RhCcEe protein show the critical amino acids, which can distinguish between the (Rh C and Rhc) / (RhE and Rhe).

Adapted from Avent (2018).

The role of the first histidine pore is a prerequisite in transporting the methylammonium (Javelle *et al.*, 2006). RhD and RhCE are not involve in ammonia transport but may be evolving a new function in the RBC membrane (Westhoff & Wylie, 2006).

1.6 *RH* polymorphisms

The Rh system has two related genes, the *RHD* gene that encodes the RhD protein, which expresses the D-antigen while the *RHCE* gene that encodes RhCE protein, which is carrying multiple specific (C, c, E, and e) antigens (Mouro *et al.*, 1993). All Rh system antigens arise from either RhD and RhCcEe polypeptides or hybrid molecules caused by an exchange between the *RHD* and *RHCE* genes.

There are two nomenclature systems for the Rh system: Fisher and Race developed the first one while Wiener developed the second one. Both systems were named after their discoverers, and they reflect the inheritance theories (Fisher & Race, 1944; Wiener 1949).

The Fisher-Race system is more frequently in use in blood banking, it is based on CDE nomenclature while the Wiener system uses the R, r nomenclature (Table 1.2).

Figure has been removed due to Copyright restrictions.

Table 1.2. The haplotype nomenclature of *RHD* and *RHCE* genes:

According to The Fisher-Race and Wiener systems and it shows the D positive nomenclature is shown in the first 4 lines and the D negative nomenclature is shown in the rest. Rosenfield *et al.*, (1973)

The primary antigens in the Rh system are D, C or c, and E or e. They were written in alphabetical order (C, D, E) but recently they have changed the order to become (D, C, E) according to Fisher-Race system to reflect gene order. The *RHD*, *RHCE* haplotypes are shown in Table 1.2 with the corresponding letter from Wiener system. d is used to represent the D-negative genotype. They are also referred as R and r depending on D-antigen case if it is expressed or not, respectively (Rosenfield *et al.*, 1973) (Table 1.2).

1.6.1 D-epitopes Models

There are two different predicting models for the D-epitopes called single footprint and the D epitope cluster model (Figure 1.7). In 1998, Chang and Siegel suggested the single footprint model to study the Rh (D) epitopes and the Rh D antibodies. They propose a model that reconciles the serological diversity between the anti-Rh (D) antibodies and the topological constraints imposed by the Rh (D) antigen. Their data were proven that there are a genetic homology between antibodies and D epitopes. It was based primarily on the sequence analysis of monoclonal anti-D.

Figure has been removed due to Copyright restrictions.

Figure 1.7. Predictive models of D epitope configuration.

This figure is comparing between the D epitope cluster model by Liu and the single footprints model by Chang and Siegel. In the first model it shows the external loops by cross-hatched rectangles, numbered circles indicate 12 domains and intracellular loops by black rectangles. This model is based on expression mutant's Rh cE proteins with different combinations of *RHD* loops recombinant (3 + 4; 4 + 6, 6, and 3 + 4 + 6), which was generated by site-directed mutagenesis. Studying the D epitope they used specific antibodies, the results was that the loop 4 are the critical loop. The single footprint model of D epitope, it shows the external loops by cross-hatched rectangles, 12 domains are shown by numbered circles and intracellular loops by black rectangles. This model argued that D epitopes are close and similar to each other and the only difference is the number and arrangement of the amino acid residues that can be D-specific, which shows as black circles or D/CcEe common protein residues, which shows as white circles.

Adapted from Avent (1999).

Moreover, Rh (D) epitopes classically defined by panels of Rh (D) variant cells are not spatially discrete. The data showed that there is a wide genetic homology between antibodies against different Rh (D) epitopes. In spite of the specificity differences of the RhD epitopes, these antibodies, which share a high degree of structural relatedness and have the ability to inhibit the binding between each one of them (Chang & Siegel, 1998).

Furthermore, in 1999, Liu *et al.* suggested the second model, which is the D epitope cluster model. In this study, they suggest 6 clusters model for the D epitope. This study was based on replacing the external loops of RhCE by the external loops of RhD (loop3, loop4, and loop6) they have used each loop individually and then all loops simultaneously. At the end, they construct RhCE with external loops (3, 4, and 6) of RhD. They used the Site-directed mutagenesis (SDM) technique. A human myelogenous leukaemia cell line K562 cell was used to see the constructed gene expression; where *RHD*-derived amino acids were expressed on chimeric RhCcEe cDNAs.

The purpose of these constructs was to study the expression of individual D epitopes, which led to the prediction of the external loop-loop interactions of the RhD protein and predicted the molecular topology of Rh antigen epitopes. Flow-cytometry was used with a total of 50 MAb anti-D to analyse the different expression pattern of D epitopes (Liu *et al.*, 1999 a; Liu *et al.*, 1999 b).

Each model was claiming about different points: the single footprints model was arguing that D epitopes are similar, and the only difference is about the number and the arrangement of the amino-acid in the paratope-epitope contact area or the common protein residues. Moreover, they believed that any RhD antibody could bind to the different D epitope at the same time regardless of the distance between these epitopes. Chang & Siegel

moreover did not sequence antibodies that react with the 6th exofacial loop of the RhD protein. While the D epitope cluster model was arguing about the distance between each loop, which must be at least 20Å apart, and therefore, the Rh D protein monomer is more than 50Å in diameter.

According to the X-ray crystallography of the antibody-antigen interaction it shows that the approximate diameter of the paratope is 20 Å. Commonly the paratopes make contact with a minimum of 15 to 20 amino acids in the corresponding epitope (Davies & Cohen, 1996). The distance between each external loop of the RBCs membrane protein is around 20-30 Å (Walz *et al.*, 1997; Cheng *et al.*, 1997).

Thus, the paratope of the antibody is capable of contacting with up to 3 external loops. Therefore, that lead to disagree with the footprint model because it is based on using the one antibody to react with all the different epitopes at the same time, which is impossible due to the distance. So, these models are conflicting with each other in the manner of the distance between each epitope and the number of epitopes that anti-D can bind to at any one time.

1.7 D-Variant Phenotypes

Although the phenotypes of *RHD* in most people are either D+ or D-, there are other two types of D-variants, which are known

as weak D and partial D.

There are two systems to classify the D-epitopes; one uses a nine-epitope model, and the other represents 30 different epitopes. Anti-D is present in the sera of D-positive individuals. Therefore, by using the same anti-D in the D-negative people and comparing the reactivity of the D-negative to D-positive sera, it was possible to determine the different D-variants (D II to DVII) (Tippett & Sanger, 1962, 1977).

Using a monoclonal anti RhD, demonstrated the variety of RhD epitopes due to the alteration of the RhD protein structure and defined those lacking on partial D phenotype red cells. The phenotypes of the partial D are DIIIb, DIIc, D1Vb, D Va, D VI, DFR, and DBT (Flegel *et al.*, 1996), but over 100 have now been described (<http://www.rhesusbase.info/>).

A loss of one or more of the nine D epitopes can lead to Partial-D phenotypes (Lomas *et al.*, 1989, 1993). By discovering the other classification system for the D-epitope (30 epitopes system) the reason of having partial D-epitope was changed to be related to the loss of *RHD* exon and replaced it by *RHCE* exon (Jones *et al.*, 1995) or point mutations in *RHD* affecting externalised regions of the molecule.

DNU and DII phenotypes are based on a single mutation in exon 7 on the *RHD* gene. DNU phenotype the amino acid residue number 353 on the RhD sequence substitute from Gly to Arg while in the DII phenotype an amino acid change in the RhD se-

quence in position 354 from Ala to Asp, which is responsible for the lacking of epD4 and epD9 (Avent *et al.*, 1997). (Rhesus base website, <http://www.uni-ulm.de/fwagner/RH/RB><http://www.uni-ulm.de/fwagner/RH/RB>) provides an up to date resource of all known partial and weak D phenotypes.

1.8 Rh in Various Species

Mammals appear the only species able to express RhD and RhCE proteins, while the RhAG- related protein seems to be present in different species, including marine sponges, insects, nematodes, slime moulds and fish (Huang & Liu, 2001). RhAG protein in RBCs shows to accelerate ammonium transport (Ripoche *et al.*, 2004). Furthermore, it has the same function in yeast (Marini *et al.*, 2000; Westhoff *et al.*, 2004). Moreover, in *Xenopus* oocytes RhBG and RhCG both work to facilitate the ammonium transport (Bakouh *et al.*, 2004; Ludewig, 2004).

The RhD and RhCE proteins are different functionally to RhAG, as they don't transport ammonium (Westhoff *et al.*, 2002; Ripoche *et al.*, 2004).

The RhAG protein shows a similar structure and a similar three-dimension structure of other Rh proteins, and that's attributable to the trans-membrane helix (TMH) regions (Liu *et al.*, 2000; Conroy *et al.*, 2004; Khademi *et al.*, 2004; Zheng *et al.*, 2004). Moreover, the RhD and RhCE proteins are identical to each

other with a tiny difference. The RhCG molecule (Figure 1.8) has been crystallised and shows similar topology to that predicted by Conroy *et al.*, 2005. Furthermore, Rh proteins show sequence homology to ammonium transport proteins (Marini *et al.*, 1997).

According to Conroy 2005, predicted topology model of human RhD and RhAG proteins (Figure 1.5) and *Escherichia coli* (*E.coli*) ammonia channel AmtB model (Figure 1.9) (Khademi *et al.*, 2004; Zheng *et al.*, 2004) with multiple sequence alignments of the unknown Rh family member. Also, using of the *ClustalW2* Multiple Sequence Alignment

(<http://www.ebi.ac.uk/Tools/msa/clustalw2/>) (Thompson *et al.*, 1997) as an alignment tool, Avent (unpublished data, 2015) aligned the 13 Rh proteins, including human RhD, RhCE, RhAG, RhCG, Murine RhAG, Murine RhBG, Murine RhCG, *Geodia cydonium*, *Caenorhabditis elegans* Rh 1, *Caenorhabditis elegans* Rh 2, and *Nitrosomonas europaea* NeRh50. Their accession numbers from the GenBank are: (*ALM96707*, *ANT95300*, *AH-Y04440*, *NP_057405*, *NP_035399.1*, *NP_067350*, *NP_062773.2*, *C-AA73029*, *AAF97864.1*, *AAF97865.1* and *WP_011111083*) (Figure 1.10). These elements combined help in mapping the sequence of interest in this project.

All known Rh-related proteins from bacteria, plants, insects and animals that share an equivalent structure of the conserved regions, which are located in the central subunit pore. The most

important conserved regions are The Phenylalanine Gate, The Twin-His Motif, and ammonium transporter fold.

Figure has been removed due to Copyright restrictions.

Figure 1.8. Secondary structure of RhCG protein.

Adapted from Gruwitz *et al.*, (2010).

Figure has been removed due to Copyright restrictions.

Figure 1.9. Crystal structure of AmtB:

Adapted from Laganowsky *et al.*, (2014).

MRFTFPLMAIVLEIAMIVLFGFVEETQTVLEQLNITKPTDMGI-----	46 Hum RhAG
MAWNTNLRWRPLPTCLLQVIMVILFGVFVRYDFAADAHWWSERTHKNLSDMENE-----	55 Hum RhCG
MREKFSLIALSLEVVMIVSFALFVEYETSQNGSQKSAPQONASQQNAAQQNASQQGNAS	60 <i>R.norvegicus</i>
MREKFPPLMAISLEVAMIVLFGFVEETPQNASQKNASHQNASQQGNTSSSA-----	52 Murine RhAG
MARVPRHRLVLPILCLLFQGATALLFAIFVRYNHETDAALWHWGNHNSV-----	50 Murine RhBG
MAWNTNLGRIPITCLIQVIMVVLFGVFVRDIQADAHWWLEKKRKNISSDV-----	53 Murine RhCG
MDWAKMLLPGLLVFQVFIILYGLLVRYDDTGDAIRNDTTISDVSNLDSYRST-----	54 <i>G.cydonium</i>
MRSPLHQNLTIILGLFQVFLVIFALYGSY--DAS-ALPSETKNVEEAAR-----	48 <i>C.elegans Rh1</i>
MWSVLHRRQFAIIAGLMQTVFIVLFAKYVKYIDPLDSRRVYSGTD-----	46 <i>C.elegans Rh2</i>
MHSPAARKVSGYVLMIVQIIFLVLFVLFVREKLTALPLA-IDAEDAGSANE-----	50 Drosophila
MSSKYPRSVRCPLTALTLEAALILLFYFFTHY--DAS-LE-----DQKG-----	43 Hum Rh Ce
MSSKYPRSVRCPLTALTLEAALILLFYFFTHY--DAS-LE-----DQKG-----	43 Hum Rh D
MSKHLCTAFSSIALFLLCFSSWASAVAPAEINEAR	36 NeRh50
-----FFEYPLFDVHVMIFVGFGLMTFLKRYGFSVSGINLLVAALGLQGTIVQGIL-Q--SQGQ	106 Hum RhAG
-----FYRYPSFDVHVMVFGFGLMTFLQRYGFSVAVGFNLLAAGFIQWALLMQGWFFH--LQDR	116 Hum RhCG
SPAKEDQFFQLYPLFDVHVMIFVGFGLMTFLKRYGFSVGVGFNLLAALGLQGTIVQGIL-H--SHGL	127 <i>R.norvegicus</i>
---KKDQFFQLYPLFDVHVMIFVGFGLMTFLKRYGFSVGVGFNLLAALGLQGTIVQGIL-H--SHGK	116 Murine RhAG
---DNEFFRYPSFDVHVMVFGFGLMTFLQRYGFSVGVGFNLLAALGLQGTIVQGIL-H--SFHG	113 Murine RhBG
---ENEFFRYPSFDVHVMVFGFGLMTFLQRYGFSVGVGFNLLAAGFIQWALLMQGWFF-HYFEEG-	117 Murine RhCG
-----L-KVMPFFDQVHVMIFVGFGLMTFLKRYGFSVSGINLLASFAIQWSTLTSGVF-Q--FIDQ	113 <i>G.cydonium</i>
-----MTNLYPLFDTHVMIFIGFGLMTFLKRYGFSVSNMLLAVFTIQWGIIVRGMASA--HHGF	109 <i>C.elegans Rh1</i>
-----YPLFDVHLMIFVGFGLMAFLKRYGFSVSNLLSAFVIOFAMLRGFMT--VAFQ	102 <i>C.elegans Rh2</i>
-----HVSXYPQFDIQVMIFIGFGLMTFLKRYGYSATGFIIFMAALVVOAVLMKGFH--MEGG	110 Drosophila
-----LVASYQVGDLTVMAALGLGLTSTNFRHSWSSVAFNLFMLALGVQWAILLDGFLSQ--FPPG	104 Hum Rh ce
-----LVASYQVGDLTVMAALGLGLTSTNFRHSWSSVAFNLFMLALGVQWAILLDGFLSQ--FPPG	104 Hum Rh D
LVAQNYNSINILALLVGFGLMVFVRVYGSATTGTYLVVATGLPLYILL-RANGI--FGHA	96 NeRh50

166 Hum RhAG
 176 Hum RhCG
 187 *R.norvegicus*
 176 Marine RhAG
 174 Marine RhBG
 177 Marine RhCG
 180 *G.cydonium*
 169 *C.elegans* Rh1
 165 *C.elegans* Rh2
 170 *Drosophila*
 163 Hum Rh ce
 163 Hum Rh D
 156 NeRh50

 232 Hum RhAG
 242 Hum RhCG
 253 *R.norvegicus*
 242 Marine RhAG
 240 Marine RhBG
 243 Marine RhCG
 244 *G.cydonium*
 234 *C.elegans* 1
 230 *C.elegans* 2
 237 *Drosophila*
 230 Hum Rh ce
 230 Hum RhD
 227 NeRh50

 299 Hum RhAG
 309 Hum RhCG
 320 *R.norvegicus*
 309 Marine RhAG
 307 Marine RhBG
 310 Marine RhCG
 312 *G.cydonium*
 301 *C.elegans* Rh1
 297 *C.elegans* Rh2
 320 *Drosophila*
 297 Hum Rh ce
 297 Hum Rh D
 292 NeRh50

 360 Hum RhAG
 370 Hum RhCG
 381 *R.norvegicus*
 370 Marine RhAG
 368 Marine RhBG
 371 Marine RhCG
 378 *G.cydonium*
 369 *C.elegans* Rh1
 365 *C.elegans* Rh2
 388 *Drosophila*
 361 Hum Rh ce
 361 Hum Rh D
 356 NeRh50



Figure 1.10. Sequence alignment of the Rh protein family members.

A total of 13 proteins from the Rh family is aligned using *ClustalW2 Multiple Sequence Alignment* (Thompson *et al.*, 1997). The green highlighted amino acids are referring to a Consensus Rh protein family Homology. The purple highlighted amino acids are referring to Human Rh glycoprotein vs. *G. Cydonium*/*C. Elegans* homology. The cyan highlighted amino acids are referring to Human Rh glycoprotein Vs. Human Rh polypeptide homology. The yellow highlighted amino acids are referring to differences, human Rhce vs. Rhd polypeptides.

Adapted from Avent (2015).

1.8.1 *Escherichia coli* (*E.coli*)

In 1885, *Escherichia coli* (*E. coli*) was discovered in a healthy individual's feces by a German-Austrian paediatrician called Theodor Escherich, who firstly named it as *Bacterium coli* refer to presence place (colon) and he was able to determine the morphology features and the properties of this bacteria. Later in 1919, after Theodor Escherich passed a way the bacteria name was changed to reflect Theodor Escherich name.

E. coli is a rod shaped, anaerobic, Gram-negative bacterium, which is naturally found in the human body. Although most *E. coli* strains are harmless there are some harmful strains, which can cause a food poisoning to the host body. *E. coli* has a benefit to the human host by producing vitamin K2 (Bentley & Meganathan, 1982). Due to the simplicity of using, growing *E. coli* in the laboratory, it consider to use widely as a model or a host organism in biotechnology and microbiology specially when work with recombinant DNA (Lee,1996; Makrides 1996).

E. coli is surrounded with lipopolysaccharide (LPS) layer, which forms the cell wall membrane. *E. coli* also surrounded with an outer membrane and work as a selective barrier to provide the *E. coli* with high protection against antibiotics such as penicillin (Raetz,1996; Gerard, 2010).

E. coli is reliant on ammonium as a source for nitrogen (Kleiner, 1985) via ammonium transporter (AMT) family, which are integral membrane proteins (Von & Merrick, 2004). The *E. coli*

AmtB was the first Amt protein to be purified (Blakey, 2002). It has a stable homotrimer in the cytoplasmic membrane. Topology of *E. coli*'s AmtB shows to have 11 transmembrane domains with 5 external loops N-terminus out and C-terminus inside the cytoplasm (Thomas *et al.*, 2000; Berneche *et al.*, 2004; Khademi *et al.*, 2004; Scanlon & Merrick, 2006). (Figure 1.11)

Figure has been removed due to Copyright restrictions.

Figure 1.11. Topology structure of *E. coli* AmtB Channel:

This figure shows the AmtB amino acid sequence arrangement within the membrane. It has 11 membrane domains label as M1-M11. These domains bind together to create 5 external loops facing preplasm side. Red circle residues contribute to the substrate contacting walls of the channel. Blue circle residues contribute to the inter-monomer contacts.

Adapted from Khademi *et al.*, 2004

1.8.2 *Nitrosomonas europaea*

Nitrosomonas europaea (*N. europaea*) are Gram-negative bacterium; their genome is a single circular chromosome of 2,812,094 bp. These bacteria exists in several places such as; soil, sewage (McTavish *et al.*, 1993). They can derive their energy from the oxidation of ammonia to nitrite by using a hydroxylamine oxidoreductase enzyme, which catalyses the oxidation of hydroxylamine to nitrite (Arp *et al.*, 2002).

Studies of the expression of the *NeRh50* protein on *Saccharomyces cerevisiae* Δ *mep* bacteria strain proved the function of *NeRh50* protein (Figure 1.12). In this study, the Rh50 protein of *N. europaea* and *Nitrospira multiformis* (*N. multiformis*) was used to replace the lack of *MEP* gene in the *Saccharomyces cerevisiae* Δ *mep*, and to restore the NH_4^+ uptake, which is critical to bacterial growth. That indicates *NeRh50* proteins can replace Amt proteins to aid ammonium uptake. This process is not presently well understood, although Rh50 could potentially play a role (Marini *et al.*, 2000; Cherif-Zahar *et al.*, 2007; Weidinger *et al.*, 2007). Moreover, there were many studies done to explore the Rh proteins possible role in CO_2 transport, but the results of these studies were impeded by the lack of a reliable biochemical assays (Peng & Huang 2006).

According to the topology prediction of the Rh50 like protein, it shows that it has 12 membrane domains. The 1.85-Å resolution was used to determined this structure. Thus, choosing

this kind of bacteria to construct the hybrid recombinant gene was the similarity between the human Rh protein and the *N. europaea* in the topology aspect. Both proteins have 12 membrane domains.

Figure has been removed due to Copyright restrictions.

Figure 1.12. The trimer structure of the Rh-like protein in *Nitrosomonas europaea* (*NeRh50*).

The yellow circle in the left monomer face surface structure shows the extracellular pore entry. At the right, ribbon representation it shows the side view.

Adapted from Lupo *et al.*, (2007).

Figure has been removed due to Copyright restrictions.

Figure 1.13. Topology structure of *NeRh50* protein in *N. europaea*:

1.9 The Immunogenicity of Rh Blood Group

Due to the clinical significance of an incompatible blood transfusion, which can lead to severe or even fatal cases of damaging organs and cells specifically RBCs, the importance of knowing the correct compatible blood and how to treat patients in case of any blood group mismatching was raised.

Transfusion reactions are caused by antibodies against blood group antigens. These antibodies may be naturally existing; for example, antibodies, mostly IgM, against the absent antigens of the ABO system are naturally present in blood. On the other hand, alloimmunisation with immunoglobulin G (IgG) antibodies occurs as a consequence of the exposure to blood group antigens absent in an individual, which generally occurs via transfusion or foeto-maternal haemorrhage during pregnancy. Antibodies (Abs) against RhD in the cases of D negative persons will not appear in their plasma. Unless this person has been exposed to D positive blood type, for example, in the case of a D-negative mother with a D-positive foetus as in HDFN (Mollison *et al.*, 1987; Daniels, 2013). The Rh blood group system is clinically relevant, because antibodies against Rh antigens are involved in HDFN, HTR, and autoimmune haemolytic anaemia (Mollison *et al.*, 1987; Evers *et al.*, 1994).

1.9.1 Haemolytic Transfusion Reactions

Haemolytic transfusion reactions (HTR) are relatively rare conditions due to incompatible blood transfusion. HTR is caused by alloimmunisation, which is an immune response to any foreign antigen from different cells or tissues. So it occurs in blood transfusion and organ transplantations. HTR happens due to the incompatible blood group between the donor and recipient (Sloand *et al.*, 1994). The recipient immune system starts to attack the donor RBCs. This kind of reaction can happen during or after the transfusion.

The complications related to the transfer of an incompatible blood group are ranging from mild reactions to severe reaction due to the short lifespan of the transferred RBCs and thus lead to have high levels of bilirubin and lactate dehydrogenase. Milder reactions are; fever, chills and anaemia, while the severe reactions are; coagulation, renal failure and death (Eder & Chambers, 2007).

1.9.2 Haemolytic Disease of the Foetus and Newborn

The historical story of this disease was begun in 1609. It was firstly described by a midwife called Louise Bourgeois when she first saw the hydrops foetalis in a twin baby's case (Bowman, 1988). Then, in 1939 Levine and Stetson found an antibody

that caused HDFN (Levine & Stetson, 1939).

HDFN occurs in subsequent pregnancies of the RhD negative mother carries RhD positive foetus, which means, the foetal RBCs contain the D antigens while the mother's RBCs have not. In this case, the mother must have been previously sensitised to RhD positive blood (Pollack *et al.*, 1968; Bowman 1988). The sensitization happens when a woman with RhD negative blood is exposed to RhD positive blood, usually during a previous pregnancy during mainly delivery with a RhD positive baby. Blood exchange can occur between the foetus and the mother, which activates the mother's immune system to form antibodies against this foreign D-antigen. The woman's body responds to the D⁺ blood by producing antibodies that destroy the foreign blood cells (Urbaniak *et al.*, 2000).

The woman's immune system will have IgM against anti-D after the first sensitisation. Furthermore, if she is exposed to RhD positive blood again through a RhD positive foetus, her body will respond immediately by producing IgG antibodies, which are capable of crossing the placenta (Alberts *et al.*, 2002). The antibodies can continue attacking the baby's red blood cells for a few months after birth. The consequence of antibody crossing the placenta can cause jaundice in mild cases (Picchiassiet *al.*, 2015) and may lead to foetal death in severe cases (Bowman, 1998; Anstee *et al.*, 2010). The incidence of the mortality of the babies from this disease in the era between 1940 and 1970 has

shown a dramatic reduction, from 50% to 5–9% of all prenatal deaths (Bowman 1988). Over the next 20 years, HDFN deaths caused by immune anti-D were dropped from 46 per 100,000 pregnancies to 1.6 per 100,000 (Pilgrim *et al.*, 2009).

In 1968, there was a clinical trial by Pollack *et al.*, on the effective dose of the RhoGAM, Rh0 (D) Immune Globulin anti-D. This treatment was given to the pregnant women and administrated as intramuscular injection. In order to prevent the primary Rh-immunization of those pregnant mothers with an Rh-negative fetus, who had giving birth to ABO-compatible Rh-positive infants. The result of this clinical trial of intramuscularly administered RhoGAM in the maternal women showed the best dose of this treatment, which was 300 g and it, was given a great result against the Rh-immunization (Pollack *et al.*, 1968).

The most efficient treatment is avoidance of the initial RhD alloimmunization by prophylactic administration of passive anti-D immunoglobulin. Prophylaxis is given against Rh-immunisation using anti-D Ig for pregnant Rh-negative mothers. These mothers should have anti-D prophylaxis injection during 28 to 30 weeks of her pregnancy and at birth. This treatment does not have a life-long effect, so, the mother with *RHD* negative needs to have this injection in each pregnancy (Urbaniak *et al.*, 2000). HDFN disease is uncommon in the UK because it can usually be prevented using injections of anti-D immunoglobulin. It is recommended that all maternal mothers have a blood test in the

early stage of their pregnancy to determine their blood type. Directly before delivering the baby, there is some necessary test has been done to the fetus such as; newborn blood group. Direct anti-globulin test and checking the haemoglobin concentration, should be taken from the cord blood. Moreover, bilirubin level should be monitored in the first days of the neonatal period. All D-negative pregnant women have an antibody screen to detect anti-D and other clinically significant antibodies. Recombinant Rh antigens could readily be used in these cases.

1.10 Aims and Objectives

Currently, regular blood tests to define the blood phenotype of the patient are able to detect only the D and ABO antigens, but for identifying the other antigens further serological tests are required. Moreover, in cases such as alloimmunisation, extended serological testing to unknown antibodies is costly. To detect rare antibodies needs lots of reagents and rare red cells which may or may not be available to the serologist.

Recombinant DNA (rDNA) is based on joining DNA molecules from different sources to create a new DNA by using a molecular cloning technique. The importance of this approach is related to biotechnology, medicine and research purposes. This technique is advantageous due to the similarity in the chemical structure between all organisms (Watson 2007).

Recombinant Rh antigens may become an extensive protocol to use at the blood banks and laboratories because of the affordability, and the specificity of using bacterial cell membrane comparing to the human RBCs, which have a short shelf-life and low specificity especially in the case of multi-transfused patients. Therefore, the objective of this project was to develop Rh hybrid recombinant genes in bacteria, and it will be capable of detecting the D antigen with affordable price and much faster technique, and with unlimited supplies of antigens. Moreover, to generate the Rh-antigens in the bacterial membrane, these

cells will have no other human antigens on their surface, that means in the case of patients who have multiple antibodies we can use a mono-specific antigen to detect and to sort out which antibody they have.

The aims of this research were:

- 1- Constructing a hybrid *N. europaea*- human *RHD* gene by using the available synthetic cDNA of *N. europaea* will be the first step to achieve this goal, choosing the *E. coli* membrane to study the expression of the human RhD external loops (2, 3, 4, and 6), due to the simplicity of manipulating *E. coli* in the laboratory.
- 2- Cloning of these recombinant genes in the appropriate host will be confirmed by using several techniques such as synthetic cDNA, cloning, flow-cytometry and Western blotting to detect bacterial proteins. Moreover, to detect the expressing RhD protein in the *E. coli* membrane, specific antibodies will be obtained.
- 3- Generate a new, fast, and affordable diagnostic assay based on hybrid recombinant *N. europaea*- human *RHD* gene.
- 4- Demonstrate that the recombinant RhD protein can mimic the Rh blood group antigen (D) and be used to screen human sera for their antibodies. This diagnostic assay will be capable of detecting anti-D and other Rh antibodies in patient serum. By using human RhD loops (6, 4, 3 and 2), which may be expressed in the *E. coli* membrane and if successful in generating

D-epitopes, further recombinant Rh antigens may be produced.

5- This *E. coli* may be commercially available and could replace the use of red blood cells to detect patient antibodies. This *E. coli* will express just one Rh antigen and no others, although bacterial carbohydrate structures may mimic red cell antigens.

Chapter 2

Materials and Methods

2.1 Materials

2.1.1 Primary Antibodies

Tissue culture supernatant monoclonal antibody (mAb) anti-D 2.2B9, raised against a synthetic peptide corresponding to the fourth external domain of the human RhD protein (Skinner, Musherns & Avent, unpublished) and it was produced in conjunction with the International Blood Group Reference Laboratory (IBGRL), NHSBT, Bristol, U.K.

Tissue culture supernatant mAb anti-D MS-26, raised against a synthetic peptide corresponding to EpD23 of the RhD protein (Scott M, 1996). It was purchased from Plasmatec Company, Bridport, U.K.

Several purified anti-D mAb were purchased from IBGRL, NHSBT, Bristol, U.K.: BIRMA-D6, BIRMA-D10, BRAD-3, HAM-A, RUM-1, and MAD-2.

MAb anti-*E. coli* DH5 α (HRP) was purchased from Biorbyt Company, Cambridge, U.K. (Table 2.1).

2.1.2 Custom Primary Antibodies

Several antibodies were customized in our lab and produced by Dundee Cell Products, Dundee, U.K.: mAb anti-D (2A4/5G9), raised against a synthetic peptide corresponding to the sixth external domain (NH₂-CLVLDTVGAGNGMC-COOH) of the human RhD protein.

MAb Anti-*NeRh50* (1917), raised against a synthetic peptide corresponding to the first external domain (CSSWASAVAPAEINEAR) of the *NeRh50* wild type.

Polyclonal antibody anti-*NeRh50* (1918), raised against synthetic peptides corresponding to the first external domain (CSSWASAVAPAEINEAR) and C-terminus (CEFIHLAGPED) domain of the *NeRh50* wild type protein.

For producing these antibodies there was a two phase procedure. Phase one included: synthesis of stable and water soluble peptides that were coupled to carrier proteins (KLH and BSA) Following this the quality of the peptide sequence was confirmed by mass spectrometry and HPLC analyses.

The second phase included: immunisation of four mice for six to eight weeks with 3-5 injections of 50 μ g antigen based on their response to this antigen. After the cell-fusion of splenocytes and myeloma cells, which took around 9 weeks a screening test was performed for hybridomas to select the positive clones. Then, in week 11 propagation of hybridomas, cryopreservation and production of cell bank took place. Between weeks 11-13 positive clones were recloned and the Ig isotypes were characterized. Finally, in weeks 14 and 16 a cell bank and 25 ml of supernatant from each positive clone were produced (Table 2.1).

Antibody	M / P	Ig subclasses	Target Protein
2.2B9	M	IgG	Human RhD Loop 4
2A4	M	IgM, kappa	Human RhD Loop 6
5G9	M	IgM, kappa	Human RhD Loop 6
BIRMA-D6	M	IgG1	Human RhD Loops 4, 6
MS-26	M	IgG	Human RhD Loops 3, 4, 6
BIRMA-D10	M	IgM	Human RhD Loops 3, 4, 6
RUM-1	M	IgM	Human RhD Loops 3, 4, 6
HAM-A	M	IgM	Human RhD Loops 3, 4, 6
MAD-2	M	IgM	Human RhD Loops 3, 4, 6
BRAD-3	M	IgG3	Human RhD Loops 3, 4, 6
1917	M	IgG	NeRh50 Loop1
1918	P	IgG	NeRh50 Loop1, C-Terminus

Table 2.1. Polyclonal(P)/ monoclonal(M) Anti-D and Anti-*NeRh50* antibodies

List shows polyclonal(P)/ monoclonal(M) antibodies with the target protein in the Human RhD external loops and the target proteins in the *NeRh50* and their Ig subclasses, which were used in western blotting, flow cytometry and TEM techniques.

2.1.3 Secondary Antibodies

Horseradish peroxidase-conjugated rabbit anti-mouse secondary polyclonal antibody (1:1000) and horseradish peroxidase-conjugated swine anti-rabbit secondary polyclonal antibody (1:1000) catalogue numbers (P 0260; P 0399), respectively were purchased from Dako Agilent Pathology Solutions, U.K.

Goat Anti-Mouse IgG H&L (FITC)(ab6785) and Goat Anti-Rabbit IgG H&L (FITC) (ab6717) were purchased from Abcam, U.K. FITC Mouse Anti-Human CD43 1G10 (RUO) was purchased from BD Biosciences, U.K.

Goat anti-Mouse IgG (whole molecule)–Gold antibody and Goat anti-Rabbit IgG (whole molecule)–Gold antibody were purchased from Sigma-Aldrich, Gillingham, U.K.

2.1.4 Plasmids and Bacterial Strains

2.1.4.1 Plasmids and Expression vectors

pUC57 (GenScript, UK) was used as a carrier vector for the recombinant genes and was used for the initial cloning. Plasmids pESV2 was constructed at Department of Molecular Microbiology, John Innes Centre, Colney Lane, Norwich NR4 7UH, United Kingdom by Mike Merrick *et al.*, (Cherif-Zahar *et al.*, 2007). pET-9a DNA - Novagen (Merck Millipore, UK) used for expressing the recombinant gene constructs.

2.1.4.2 Bacterial Strains

E. coli DH5 α competent cells (Thermo Fisher Scientific, Paisley, U.K.) were used for cloning and *BL21(DE3)* (New England Bio Labs, Hitchin, UK) were used for expressing the recombinant gene constructs.

2.1.5 Erythrocyte samples

Ethylenediaminetetraacetate acid (EDTA) tubes containing at least 5 ml of whole blood for IP experiments were obtained from random blood donors. All samples were with complete awareness of ethical donor consent form from National Health Service Blood and Transplant (NHSBT), Filton Bristol, UK. For each sample, serology information was provided. The blood was spun at 1500 rpm for 10 minutes, and the plasma was removed, and packed red blood cells were added to 1.5 mL Eppendorf tubes.

2.1.6 Producing Luria Bertani Media (LB Media)

The LB media (Sigma –Aldrich, U.K.) was produced according to the manufacture. Once dissolved the solution was autoclaved and kept at room temperature until needed. Filter sterilized ampicillin (AMP) (100 $\mu\text{g/ml}$) / kanamycin (Kn) (50 $\mu\text{g/ml}$) (Sigma-Aldrich, U.K) were added as needed.

2.1.7 Producing Luria Bertani Agar Plates (LB Agar)

LB Agar, (Sigma –Aldrich, U.K.) was prepared according to the manufacturer’s instructions. The solution was autoclaved and kept at room temperature until needed. In cases where the mixture had solidified before use it was reheated in a boiling water bath for 1 hour. Once the temperature of LB agar had reduced to 55 °C, filter sterilized AMP (100 $\mu\text{g/ml}$) / Kn (50 $\mu\text{g/ml}$) (Sigma-Aldrich, U.K) were added as needed. Working near an opened Bunsen burner, 25 mL of the LB agar containing appropriate antibiotic was added per 10cm polystyrene Petri dish (25 plates in total/ 500 mL). Each plate was covered with a lid and left to cool for around 25-30 mins. Plates were stored at 4 °C after solidified.

2.2 Methods

2.2.1 Alignment of Rh for different species

This step was started by comparison of the human RhD protein and 12 Rh proteins from different species. The amino acid sequences of 13 Rh proteins were aligned with *ClustalW2 Multiple Sequence Alignment* (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>) (Figure 1.10). This alignment was done by Avent (unpublished data, 2015); it shows the regions of high homology between these proteins, and the different regions.

2.2.2 Mapping of the Human RhD external loops and Designing of the Recombinant Gene

Moreover, the alignment helps to predict the mapping the human RhD external loops and match it with the corresponding external loops of the Rh50 like protein *Nitrosomonas europaea* (*NeRh50*) in order to construct the hybrid recombinant gene of interest.

By comparison with Conroy *et al.*, 2005, it was possible to identify each amino acid in the external loops of the human RhD protein through the predicted RhD protein model as shown in (Figure 1.5).

The aim was to create a hybrid sequence with the *NeRh50* backbone with the human RhD sixth external loop inserted. So, to do this, the codon usage table of *E. coli* (Table 3.1), Gen-

Script website (2002-2019) GenScript Codon Usage Frequency Table(chart) Tool. [online] Available at: <http://www.genscript.com/tools/codon-frequency-table> [23-06-2019] was used to translate the amino acid sequence back to nucleotide level.

The final DNA sequence was sent to GenScript Company, USA to synthesise the cDNA of the Construct recombinant gene construct of *N. europaea* RH50 with a human *RHD* loop 6 (*NeRh50"6*).

2.2.3 Cloning

2.2.3.1 Cloning *NeRh50"6* Construct in pESV2 vector

2.2.3.1.1 Polymerase Chain Reaction (PCR)

2.2.3.1.1.1 Primers and Generation of *NeRh50"6* Recombinant amplicon

The primer sequences of AD1 and AD2 (Table 2.2), were obtained from Cherif-Zahar *et al.*, (2007), these HPLC primers were purchased from Eurofins Genomics, Germany. AD1 and AD2 were used to add new enzyme restriction sites in the *NeRh50"6* sequence to simplify the ligation with the pESV2 vector:

Primer Name	Sequence (5'-3')	Direction	Restriction Sites	Purification
AD1	GCCTCGAGCCATATGGGGAGTCAACAAATG	Forward primer	<i>XhoI</i> , <i>NdeI</i>	HPLC
AD2	CGGGATCCCGCTAATGGTGATGGTGGTGATG TCCTTCTGACTTGGC	Reverse primer	<i>BamHI</i>	HPLC

Table 2.2. Primer sequences to add *NdeI* enzyme restriction site in the recombinant gene *NeRh50''6*.

These primers were obtained from Cherif-Zahar *et al.*, (2007), these primers were purchased from Eurofins Genomics. The underlined parts of the primers indicate the restriction sites. In AD1 the first part of the primer indicate the *XbaI* restriction site while the second underline section indicates the *NdeI* restriction site. In AD2, the first part of the underlying sequence of the primer (GGATCC) indicates the restriction site of *BamHI*.

2.2.3.1.2 PCR Amplification of the Recombinant Gene

Q5 Hot Start High-Fidelity 2x Master Mix (from Eurofins Genomics) was used for the 50 μ L PCR reaction, the final concentrations were 1x Q5 Hot Start High Fidelity Master Mix and 0.5 μ M forward primer and 0.5 μ M reverse primer. Then, 5 μ L of DNA and nuclease free water was added to complete the volume to 50 μ L. After that, PCR were conducted as a Veriti Thermal Cycler (Thermo Fisher Scientific, U.K.), at the following conditions; 98 °C for 30 s (initial denaturation); 35 cycles of; 98 °C for 10 s (denaturation phase); 68 °C for 30 s (annealing phase); 72 °C for 60 s. Then, the 35 cycles were then followed by 2 minutes at 72 °C (extension phase) and a 72 °C.

2.2.3.1.3 Agarose Gel Electrophoresis

PCR amplicons were analysed using agarose gel electrophoresis. Samples were run on 1% (w/v) agarose gels, which were produced by mixing 1 x TAE buffer (40 mM Tris-acetate, 1 mM EDTA [pH 8.0]) with 1X GelRed Nucleic Acid Stain (Cambridge Bioscience, Cambridge, UK). 10 μ L of the PCR Ranger 100bp and 5 μ L of The Quick-Load® 1 kb Extend DNA Ladder (New England BioLabs, U.K.) DNA Ladder was purchased from Geneflow, United Kingdom suppliers of Molecular Biology Product, UK were loaded as a marker for the gel electrophoresis run.

Electrophoresis was carried out at 120V for 80 mins in 1 xTAE buffer. Following the run, gels were removed and placed in the EC3 imaging 104 system (UVP BioImaging system, Cambridge, UK) and results were analysed using the Launch Vision WorksLS program (Chemi Doc 410).

2.2.3.2 Enzymatic Digestion

0.5 μ L of *Bam*HI and *Nde*I (Promega Corporation, Southampton, UK.) restriction enzymes were used with pESV2 vector and with the recombinant amplicons PCR product following of *NeRh50"6* with AD1/AD2 primers to create cohesive ends, which made the cloning into any pESV2 expression vector easier. The digestion step was 20 min at 37 °C. Finally, inactivation of enzymes was achieved by heating to 70 °C for 10 min.

2.2.3.3 Purification of Digested DNA

2.2.3.3.1 Desalting the Recombinant amplicon

In this step, the digested recombinant amplicons *NeRh50''6**[^] were purified without losing their DNA. one tenth volume of ammonium acetate and 2 volume 100% (v/v) cold ethanol. Samples were incubated in low-temperature condition -20°C for 30 min to precipitate the recombinant amplicons DNA. The DNA was resuspended in 40 μL nuclease-free water.

Finally, Nano-Drop 2000 UV-Vis Spectrophotometer (Thermos Scientific, UK.) was used to estimate the DNA concentration in these samples.

2.2.3.3.2 Gel Extraction of Digested pESV2 expression vector

In the previous step, the restriction enzymes (*BamHI* and *NdeI*) cut the expression vector (pESV2) to give two fragments. These were run on 1% (w/v) agarose gel and following separation, the fluorescent band of the DNA fragments of the pESV2 vector were excised from the agarose gel using a clean sharp scalpel on the EC3 imaging system (UVP BioImaging system) under UV light rays the gel slices were weighed and the DNA was purified using the QIAquick gel extraction protocol (QIAGEN, Hilden, UK) according to the manufacturer's instructions, as detailed below.

Three volumes of buffer QG were added to one volume of gel

and the mix was incubated at 50 °C for 10 mins, with vortexing every 2-3 mins until the gel was completely dissolved. Then, one gel volume of isopropanol was added to increase the concentration of the DNA. Samples were moved to QIAquick spin columns and centrifuged for 1 minute at maximum speed at room temperature. 500 μ L of buffer QG was used to remove all traces of agarose and the samples were centrifuged for 1 minute at max. speed at room temperature. 750 μ L of Buffer PE was added to the QIAquick column and centrifuged for 1 min at max. speed at room temperature. The flow through was discarded after each step. The QIAquick spin columns with a clean 1.5 mL microcentrifuge tube and the samples were eluted in 30 μ L of Buffer EB (10 mM Tris-Cl, pH8.5). Finally, the DNA samples were quantified using the NanoDrop 2000 UV-V is Spectrophotometer (Thermos Scientific, UK.) and the purified DNA of pESV2 plasmid was stored at 4 °C for later use in cloning step.

2.2.3.4 Molecular cloning into pESV2

The molecular cloning process was conducted near a Bunsen Burner. The purified digested pESV2 plasmid and purified digested recombinant amplicon *NeRh50''6*^* were used to create pESV2/ recombinant *NeRh50''6* gene. LigaFastTM Rapid DNA Ligation system (Promega Corporation, UK) was used to ligate the pESV2 vector with the amplified digested recombinant gene *NeRh50''6*^*.

The following equation was used to calculate how many nanograms of the insert should be used, and from these numbers, the volumes that should be used can be determined.

$$\text{ng of Insert} = \frac{\text{ng of vector} \times \text{Kb of Insert}}{\text{Kb of vector}} \times \frac{\text{Insert}}{\text{Vector}}$$

The recombinant PCR product *NeRh50''6*^* (insert) to plasmid DNA pESV2 (vector) size is 5Kb, and the insert size is 1349 bp. The ratio of 2:1 and 5:1 (Insert: Vector) was used in the equation above.

The reaction mixture was then gently mixed, followed by incubation for 5 minutes at room temperature.

2.2.3.5 Transforming in *DH5 α* *E. coli* competent cells

Following ligation, the plasmids were transformed in to *DH5 α* *E. coli* competent cells. This step was conducted near a Bunsen Burner. The transformation was performed by adding the following reagents into a 0.5 mL sterilized eppendorf tube, in the indicated order: 100 μ L of the thawed *DH5 α E. coli* competent cells (Thermo Fisher Scientific, UK) was mixed gently with 6 μ L of the cloning mixture. The mixture was incubated on ice for 30 minutes, then, the mixture was heat shocked at 42 °C for 45 sec and then the mixture was immediately replaced on ice for 2 min.

Next, 0.9 mL of (Super Optimal Broth) S.O.C media (provided in the kit and equilibrated to room temperature) was carefully added to the mixture, which was then placed horizontally into a shaking (170 rpm) incubator at 37 °C for 1 hour.

Finally, different volumes of the transformation mixture were spread over pre-warmed selective plates (LB plates containing 100 μ g/ml with AMP). Plates were incubated overnight at 37 °C. Next day, master plate (LB plate with AMP) was prepared by inoculating single colonies from the transformed plate. These plates were incubated at 37 °C for overnight (see results sec-

tion).

Next day, a number of single colonies from the master plate were picked by sterilized plastic disposable loops and incubated individually in tubes containing 5mL medium (LB medium containing 100 μ g/ml AMP) overnight at 37 °C in shaking (170rpm) incubator.

Then, the cultured colonies with cloning expression vector (pESV2) and a recombinant DNA (*NeRh6**) were isolated and purified using the Wizard (R) Plus SV Minipreps DNA Purification System (Promega, UK) (Figure 2.1).

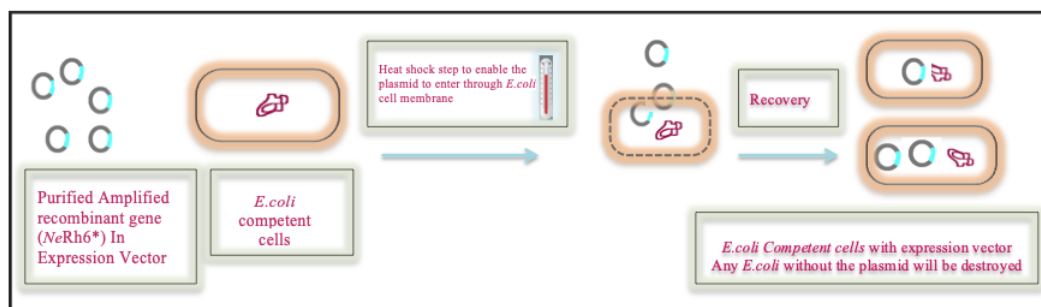


Figure 2.1. Cloning and Transformation.

Diagram shows transformation step, where the vector with the recombinant gene was transformed into the *DH5 α* *E. coli* (host cells) by the heat shock process. Cloning step, where the transformed host cells were able to replicate. Amp resistance bacteria are selected and will only grow in the presence of the *NeRh50**/pESV2 plasmid.

2.2.3.6 Isolation and purification of plasmid DNA

Wizard^(R) Plus SV Minipreps DNA Purification System (Promega,UK) was used to isolate and purify the plasmid DNA. 5 mL LB-AMP overnight broth, was centrifuged at 4500 xg for 10 mins at room temperature, the supernatant was removed gently, without disturbing the pellet, and discarded.

The pellet was resuspended by adding 250 μ L of Cell Resuspension Solution and pipetted to completely resuspend. Then, the mixture was moved to a sterilised eppendorf tube, followed by adding 250 μ L of Cell Lysis Solution and mixed by inverting the tube 4 times. To inactivate endonucleases and other proteins released during bacterial cell lysis which can adversely affect the quality of the isolated DNA, 10 μ L of Proteinase K was added to the mixture then mixed well by pipetting before incubation for 5 mins at room temperature. 350 μ L of Neutralization Solution was added and immediately mixed by inverting the tube 4 times, followed with centrifugation of the mixture at 4500 xg for 10 mins at room temperature.

Clear solution were moved (around 800 μ L each time) to the Spun Column, which was provided with the kit. The columns were spun at 4500 xg for 1 min at room temperature, the flow-through was discarded from the collection tube, and this process was continued until all the supernatant was centrifuged. Then, 750 μ L of Column Wash Solution was added, followed by centrifuging the supernatant at 4500 xg for 1 min at room

temperature, and then the flow-through was discarded from the collection tube. Then, spun column was re-washed by 250 μ L of column Wash Solution then centrifuged the spin column at 4500 xg for 2 mins at RT, the flow-through was discarded.

The Spin Column was transferred to a new, sterile 1.5 mL micro-centrifuge tube, followed by elution of the plasmid DNA by adding 30 μ L of Nuclease-Free Water to the Spin Column. The plasmid DNA was eluted by spinning at maximum speed for 1 minute at room temperature.

Finally, the eluted sample was stored at -20°C until required. 2 μ L purified plasmid DNA was quantified using the Qubit[®]2.0 Fluorometer and the Qubit[®] dsDNA Broad range (BR) assay Kit (Life Technologies, UK), according to the manufacturer's instructions.

During this procedure, 500 μ L of LB-culture broth was stored together with 50% (v/v) sterile glycerol at -80°C for future use.

2.2.3.7 Confirmation tests

2.2.3.7.1 PCR Screening

After 16 hours of incubation for the two different ratios of the ligation step (2:1 and 5:1), PCR screening was done to confirm if the recombinant gene (*NeRh50''6*^*) was inserted correctly in the expression plasmid without contamination. Steps in (See 2.2.2.3.1) were followed with AD1 forward primer, and AD2 re-

verse primer.

In 0.2 mL sterilised nuclease free tubes (BIOplastics, Landgraaf, the Netherlands) 50 μ L of PCR reaction mix were added. Using sterilized filter tips (Fisher Scientific, Loughborough, UK) single colonies from the transformation plate were picked and inoculated on a selective LB Agar master plate, which was divided by numbers (each number indicating a different single colony). Then, the same tip was moved into a PCR tube, which contained PCR master mix and was gently mixed, followed by brief spinning prior to thermo-cycling (Figure 2.2).

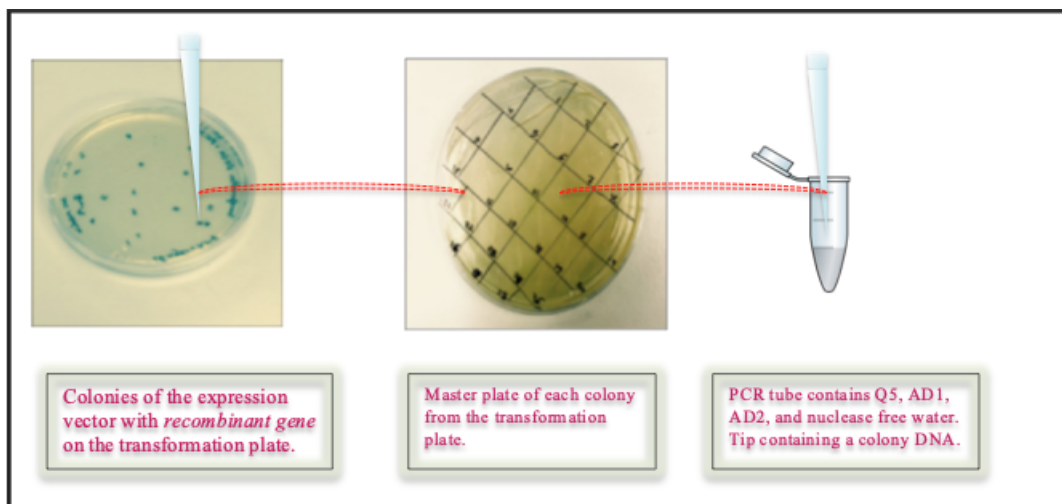


Figure 2.2. Screening process.

Using a sterilised tip to pick a single colony from the transformation plate (example colony number 11) move the tip to a divided agar plate with AMP and do a small line with each tip alone (inoculate). Then move the same tip to PCR tube, which contains the mixture of Q5 -Hot Start High-Fidelity 2x master mix, primers (AD1 forward and AD2 reverse), Use the tip just once with each colony.

The PCR tubes were run on Veriti Thermal Cycler (Thermo Fisher Scientific, U.K.) according to the conditions in (2.2.3.1.2). Gel visualisation was made by running these samples on 1% (w/v) Agarose gels to check the size of the PCR amplicons.

2.2.3.7.2 Sequencing of pESV2 gene by Next Generation Sequencing

Ion XpressTM Plus gDNA Fragment Library Preparation (Thermo-fisher scientific, UK) was used for sequencing. In a 1.5 mL Eppendorf LoBindTM Tube the following reagents were added in the indicated order 100 ng of purified genomic plasmid DNA from both the original sample (pESV2-G1000) and the pESV2- (*NeRh50"6**^), 5 μ L of Ion ShearTM Plus 10X Reaction Buffer followed by adding Nuclease-free Water to increase the volume up to 40 μ L. The mixture was vortexed for 5 sec at room temperature.

10 μ L of Ion ShearTM Plus Enzyme Mix II was added to the sample and immediately mixed by pipetting up and down. To fragment the DNA in 450 bp, samples were incubated in a heat block at 37 °C for 5 mins and immediately following this 5 μ L of Ion ShearTM Stop Buffer was added, and mixed thoroughly by vortex. The reaction tube was stored on ice.

To purify the fragmented DNA, 99 μ L of AgencourtTM AMPureTM XP Reagent was added to the sample, then the mixture of sample and beads was mixed by pipetting up and down 5 times, then the mixture was incubated at room temperature for 5 mins. Tubes were spun and placed in a magnetic rack (Thermo Fisher Scientific, Paisley, UK) for 3 mins until the solution was clear of brown tint when viewed at an angle. The supernatant were removed and discarded carefully without disturbing the bead

pellet.

0.5 mL of 70% ethanol was freshly prepared and added to the tube whilst keeping the tube in the magnetic rack followed by incubation for 30 seconds and turning the tube around twice in the magnetic rack to move the beads around. Then, the supernatant was discarded without disturbing the pellet. The wash step was repeated twice.

The tube was spun and placed back in the magnetic rack and any remaining supernatant was carefully removed this was to ensure all ethanol was removed without disturbing the pellet. The beads were air dried at room temperature for 5 mins whilst keeping the tube in the magnetic rack. Then, the tube was moved from the magnetic rack, 25 μ L of Low TE was added directly to the pellet to disperse the beads and to elute the DNA. The suspension was mixed thoroughly by pipetting up and down 5 times, and then the sample was vortexed for 10 seconds. After that, the tube was pulse-spun and placed back in the magnetic rack until the solution was clear. The clear supernatant was moved into a sterilized 0.2 mL PCR tube, without disturbing the pellet. AgilentTM 2100 BioanalyzerTM instrument and AgilentTM High Sensitivity DNA Kit were used to check the fragment size.

Next step was to ligate the adapters and to purify the ligated DNA. In a 0.2 mL sterilized PCR tube, the DNA and other

reagents were added according to Table 2.3 followed by pipetting up and down.

Component	Volume by Input gDNA 50-100 ng[
DNA	~25 μ L
10X Ligase Buffer	10 μ L
Ion P1 Adapter	2 μ L
Ion Xpress™ Barcode	2 μ L
dNTP Mix	2 μ L
Nuclease-free Water	49 μ L
DNA Ligase	2 μ L
Nick Repair Polymerase	8 μ L
Total	100 μ L

Table 2.3. Reaction Setup for Barcoded Libraries

The tube was placed in a thermal cycler and run following these conditions; 25 °C for 15 min, 72 °C for 5 min then 4 °C.

The mixture was moved to a new, sterilized 1.5-mL Eppendorf LoBind™ to achieve cleaning for the fragmented DNA, 120 μ L of Agencourt™ AMPure™ XP Reagent was added to the mixture followed by pipetting the mixture of fragmented DNA and beads up and down 5 times.

Afterward, the tube was pulse-spun and incubated for 5 minutes at RT. The tube was placed in the magnetic rack until the solution was clear, the supernatant was carefully removed and discarded without disturbing the pellet.

0.5 mL of freshly prepared 70% ethanol was added to the tube whilst the tube was still in the magnetic rack followed by incubating the mixture for 30 seconds at RT with moving the tube in

the magnetic rack. Next the clear solution was discarded carefully without disturbing the pellet. This step was repeated for a second wash.

The beads in the tube were air dried by leaving the tube in the magnetic rack at RT for 5 mins. 20 μ L of Low TE was added directly to resuspend the pellet and to elute the DNA, then, the mixture was mixed thoroughly by pipetting the suspension followed by vortexing the sample for 10 seconds. The tube was removed from the magnetic rack and pulse-spun then replaced in the magnetic rack until the solution was clear.

Clear supernatant was transferred to a new sterilized 1.5-mL Eppendorf LoBindrTM Tube without disturbing the pellet.

Pippin PrepTM instrument manual for 100–300-base-read libraries was followed to run the size-select the library step.

To purify the size-selected DNA, 90 μ L of AgencourtTM AMPureTM XP Reagent was added to the mixture followed by pipetting the mixture of fragmented DNA and beads up and down 5 times. Afterward, the tube was pulse-spun and incubated for 5 minutes at RT. Tube was placed in the magnetic rack until the solution was clear, the supernatant was carefully removed and discarded without disturbing the pellet.

Add 0.5 mL of freshly prepared 70% ethanol while the tube still in the magnetic rack followed with the same steps were done before. Air dry beads in the tube by leaving the tube in the magnetic rack at RT for 5 mins.

After that, the tube was removed from the magnetic rack and 0.5 mL of PlatinumTM PCR SuperMix High Fidelity (black cap) and 2 μ L of EqualizerTM Primers were added to the bead pellet followed by pipetting the mixture up and down. The tube was placed back on the magnetic rack for 2 mins, then carefully 50 μ L of supernatant was transferred to a new, sterilized 0.2 mL PCR tube, without disturbing the pellet. The tube was sealed and placed in the thermal cycler, and run on the following program. Dr Michele Kiernan of the Genomics Facility Systems Biology Centre at the University of Plymouth conducted the chip loading and sequencing steps. Finally, De-novo assembly software was used to align results against bacterial genome.

2.2.3.8 Cloning *NeRh50''6* Recombinant in pET expression vector

2.2.3.8.1 Enzymatic Digestion of pUC57-Recombinant gene

0.5 μ L of *Bam*HI and *Xba*I (Promega Corporation, UK.) restriction enzymes were used to cut the recombinant gene from pUC57 vector and to confirm if we have the right hybrid recombinant gene. Moreover, these restriction enzymes were used to cut the expression vector (pET) to create cohesive ends, which make the cloning into any expression vector easier (pET). Enzyme inactivation was done for 10 min at 70 °C.

2.2.3.9 Purification of Digested DNA

2.2.3.9.1 Desalting the Digested pET Expression Vector

In this step, the digested pET expression vector was purified by adding 1/10 volume of ammonium acetate and 2 volumes of 100% (v/v) cold ethanol. The sample was incubated at -20°C for 30 min to precipitate the pET DNA. The DNA was resuspended in 40 μL nuclease-free water.

Finally, Nano-Drop 2000 UV-Vis Spectrophotometer (Thermos Scientific, UK.) was used to estimate the DNA concentration in these samples.

2.2.3.9.2 Gel Extraction of Digested Hybrid Recombinant DNA

In the previous step, the restriction enzymes (*Bam*HI and *Xba*I) cut the pUC57 vector containing the hybrid recombinant construct gene. These samples were run on a 1% (w) agarose gel. The DNA fragments for the hybrid recombinant construct gene (1275 bp) were excised from the agarose gel using a clean sharp scalpel on the EC3 imaging system (UVP BioImaging system) under UV light rays and then weighed. DNA was purified using the QIAquick gel extraction protocol (QIAGEN, Hilden, UK) according to manufacturer's instructions. 2.2.2.3.3.2 for more information about gel extraction steps.

Finally, DNA samples were quantified using the NanoDrop 2000 UV-V is Spectrophotometer (Thermos Scientific, UK.) and the

purified DNA of the hybrid recombinant construct was stored at 4 °C for later use.

2.2.3.10 Molecular cloning into the pET Expression Vector

The molecular cloning process was conducted near a Bunsen Burner. The purified digested pET expression vector and the purified digested recombinant DNA *NeRh50"6*[^] were used to create the pET/ recombinant *NeRh50"6* gene. Liga Fast[™] Rapid DNA Ligation system (Promega Corporation, UK.) was used to ligate the pET expression vector with the digested recombinant gene *NeRh50"6*[^]. Followed the same equation in 2.2.3.4.

The recombinant PCR product (insert) to plasmid DNA(vector) size is 3Kb, and the insert size is 1349 bp. The ratio of 2:1 and 5:1 (Insert: Vector) was used in the equation (2.2.3.4). The reaction mixture was then gently mixed, followed by incubation for 5 minutes at room temperature.

Using *BamHI* and *XbaI* restriction enzymes in the digestion step for the pET plasmid and constructs recombinant DNA was facilitating the ligation step.

Finally, the cloning reactants were transformed into *BL21(DE3)* *E. coli* competent cells.

2.2.3.11 Transforming in *BL21(DE3) E.coli* competent cells

This step was conducted near a Bunsen Burner. The reaction was performed by adding the following reagents into a 0.5 mL sterilized eppendorf tube, in the indicated order: 100 μ L of the thawed *BL21(DE3) E. coli* competent cells (New England Biolabs, UK) was mixed gently with 6 μ L of the cloning mixture. The mixture were incubated on ice for 30 minutes, then, the mixture was heat shocked at 42 °C for 10 sec, followed by immediately, replacing the mixture on ice for 2 min.

Next, 0.9 mL of (Super Optimal Broth) S.O.C media (provided in the kit and equilibrated to room temperature) was carefully added to the mixture, which was then placed horizontally into a shaking (170 rpm) incubator at 37 °C for 1 hour.

Finally, different volumes of the transformation mixture were spread over pre-warmed selective plates (LB plates containing 50 μ g/ml Kn). Plates were incubated overnight at 37 °C.

Next day, master plate (LB plate with Kn) was prepared by inoculating single colonies from the transformed plate. These plates were incubated at at 37 °C for overnight (see results section).

Next day, a number of single colonies from the master plate were picked by sterilized plastic disposable loops and incubated individually in tubes containing 5mL medium (LB medium contacting 100 μ g/ml Kn) overnight at 37 °C in shaking (170 rpm) incubator.

Then, the cultured colonies with cloning pET expression vector and a recombinant DNA (*NeRh50''6^*) were isolated and purified using the Wizard ^(R) Plus SV Minipreps DNA Purification System (Promega,UK).

2.2.3.12 Isolation and purification of plasmid DNA

See 2.2.3.6 section

All previous steps: See 2.2.3.8; 2.2.3.9; 2.2.3.10; and 2.2.3.11 were followed for the other constructs: *NeRh50*"4, 6, *NeRh50*"3, 4, 6, and *NeRh50*"2, 3, 4, 6.

2.2.3.13 Confirmation tests

2.2.3.13.1 Sanger Sequencing

Purified plasmid DNA samples were sent to Eurofins Genomics (Ebersberg, Germany) for sequencing. Primer sets shown in Table 2.4, which were purchased from Eurofins Genomics were used to confirm: the human RhD external loops; *NeRh50*, and the pET expression vector were these primer sets cover the entire plasmid with the recombinant gene.

Samples were diluted to 10 ng/ μ L and then 15 μ L of each diluted sample was added to 10 mM forward and reverse primers in separate serial numbers barcoded labelled tube followed by covering them with a yellow cap which was provided with Mix2Seq kit (Eurofins MWG Operon).

Results were analysed using CLC Genomics workbench 9 software for alignment against expected reference sequence.

2.2.3.13.1.1 Primer Design

Several software packages were utilised to design the oligonucleotide primers for *RhD/NeRh50/pET* expression vector. Primer 3 software (<http://primer3.ut.ee/>) was initially used to design the primers; the NCBI BLAST database (<http://www.ncbi.nlm.nih.gov/>). was used to confirm the specificity of primers; the UCSC Genome Browser database 66 (<http://genome.ucsc.edu>) was also used to analyse the specificity of the primers. In addition, visualisation of the location of the amplicons within the genome was done using the in-silico PCR (virtual PCR). Primers were designed to cover the entire plasmid and each human RhD external loop to confirm the presence of all human RhD external loops. The primers were received as lyophilised from Eurofins Genomics (Ebersberg, Germany) in a High Performance Liquid Chromatography purity (HPLC), and were re-suspended by adding nuclease-free water (Ambion®, Applied Biosystems, Thermo Fisher Scientific, USA) as indicated by the manufacturer.

Primer Label	Primer Sequences	Reverse complement	Cover	From-To	Primer Sequence Length
F1	GGAGCCACTATCGACTACGC		PET before Ne/loops	328-348	20
R1	CTGCTAACAAAGCCCGAAAG	CTTTCGGGCTTTGTTAGCAG	PET before Ne/loops	487-507	20
F2	GTTTAGCAGCATCGCGCTGTTT		Ne/ Loop2/ Loop3	541-562	22
R2	GTTGCTGATAACCATTTTCGTTT	GAACGAAATGTTATCAGCAAC	Ne/ Loop2/Loop3	945-967	22
F3	TGTACATCCTGCTGCGTTTC		Loop2/Ne	761-781	20
R3	AAACCGGCACAATAACAGC	GCTGTTTATTGTCCGTTTT	Loop2/Ne	919-939	20
F4	ATTGTGCCGGTTTATCTGCT		Loop3/Ne	926-945	19
R4	GTAAAGCTCGGCCAGAAC	TGTTCTGGCCGAGCTTTAAC	Loop3/Ne	1151-1171	20
F5	GTTCTGGCCGAGCTTTAAC		Loop4/Ne	1153-1172	20
R5	CGCCATATCCACAATGCTC	GAGCATTGTGGATATGGCG	Loop4/Ne	1291-1309	18
F6	GAAAACCATTGACAC		Ne/Loop6	1452-1467	15
R6	ATGAATTCGTGGCTGTCTC	GAGGACAGCCACGAATTCAT	Ne/Loop6	1648-1668	20
F7	CCTTCCCCATTATGATTCTT		PET after all loops	2471-2490	19
R7	AAGACCATTGATGTTGTC	GCAACAACATGAATGGTCTT	PET after all loops	3063-3082	19

Table 2.4. Primer sequences to amplify regions in pET-*NeRh50*"2, 3, 4, 6

2.2.3.13.2 Enzymatic Digestion of pET-Recombinant gene plasmids

0.5 μ L *Bam*HI and *Xba*I (Promega Corporation, UK.) restriction enzymes were used at 37°C for 30 min to cut the recombinant gene from pET vector and to confirm if we have the right hybrid recombinant gene.

Enzyme inactivation step was achieved by heated up the sample to 70°C for 10 min. Then, the samples were run on a 1% (w/v) agarose gel to check the band sizes.

2.2.3.13.3 Electron Microscopy Imaging

Most of the sample preparation steps of this test was done with the help of Plymouth Electron Microscopy Centre (PEMC). In particular Dr. Alexander Strachan and Glenn Harper are thanked for their support and assistance in this work.

10 ml overnight LB broth with 50 $\mu\text{g/ml}$ Kn at 37 °C were harvested on a 4500 xg at room temperature for 10 min, then, the supernatant was discarded. Pellets were washed with 1 ml of 1x PBS. After that, samples were centrifuged at 2100 xg at room temperature for 2 min and the supernatant was discarded. This step was repeated two times.

After that, 500 μL mixture of fixation solution (2.5% (v/v) Gluteraldehyde and 4% (v/v) Formaldehyde in 0.1 M Phosphate buffer) was added to the samples to fix bacterial cells and to stabilize the protein structure.

Due to the fact that incident electrons need to penetrate through the sample, so that transmitted electron signals would be collected for imaging, the sample thickness needs to be limited (e.g. 50 to 250 nm for biological samples). Therefore specified sample preparation is necessary before staining them with antibodies. Steps include; cutting the specimen to fit the Gatan disc grinder by using ultrasonic disk-cutter. Then, low melting point wax was used to glue the specimen to the disc in order to have a plane view imaging. Next, these Grid were heated at 120 °C to bind slices rigidly. Then, it was glued for a second

time. Followed by polishing the specimen with Silicon Carbide Grinding Paper (SiC) paper to fix the specimen on a rotating polishing machine.

The specimen thickness were adjusted by controlling knob of the grinder with 10 μm until the specimen is flush with the polishing surface followed with another glue step. Then, Precision Ion polishing System (PIPS) was used to create high-quality TEM specimens with a plane-view observation.

Then, the grid was submerged in a blocking buffer (1x PBS-Tween 20 and 5% (w/v) of bovine milk powder) were used to wash the grid membrane for an hour at room temperature. After that, the blocking buffer was washed off and the primary antibodies (2.2B9, 2A4, and 1918) were added either neat or diluted in the blocking buffer, to the grid membrane as the manufacturer's instructions recommended. Then, it was stored at 4 °C for overnight.

On the next day, three washes with 1x PBS-T has been done to the grid membrane each wash was 5 min long at room temperature. Then, secondary antibody: Goat anti-Mouse IgG (whole molecule)–Gold antibody and Goat anti-Rabbit IgG (whole mole-

cule)–Gold antibody was prepared with a ratio of (1:20) according to manufacturer's instructions. It was added to the grid membrane after discarding the 1x PBS-T solution. Next, grid membrane was incubated for 1 hour at RT followed by washing

for three times with 1x PBS-T and then one wash with 1x PBS. Each wash was 5 min long at room temperature. Finally, images from the grid membrane was taken with Transmission Electron Microscope (TEM) with a 0.38 nm resolution and Gatan Orius camera.

2.2.3.14 Protein Extraction

2.2.3.14.1 Bacterial Protein Extraction from Broth Media

This step was based on using BugBuster Master Mix protein extraction reagent (Merck Chemicals Limited) 1.5 ml of overnight broth culture was centrifuged at 45000 xg for 10 min at RT. All supernatant were discarded, then, 300 μ L of BugBuster protein extraction reagent was mixed gently, prior to incubating these pESV2-*NeRh50"6* samples for 10 min at RT in the shaking incubator. These samples were centrifuged at 45000 xg for 20 min at 4 °C. The pellets and the supernatants were stored at -20°C for quantification test and stored at -80°C for long periods.

2.2.3.14.1.1 Qubit[®] Protein quantitation assay:

To estimate the concentration of the protein in the samples, quantification Qubit[™] Protein Assay Kit (Thermo Fisher Scientific, UK) was used. First, the Qubit working solution was prepared at a ratio of 1:200. Three standards were prepared according to the manufacturer's instructions, the recommended

ratio for each standard and each sample is 1:200. Next, standards and samples were incubated at room temperature for 15 mins.

Finally, the samples were measured by using Qubit[®] Fluorometer 2.0. To calculate the concentration of the protein this equation was followed:

$$\text{conc} = \text{QF value} \times 200 / (x),$$

where (x) represent the number of μL from the sample that has been used.

2.2.3.14.2 Erythrocyte Ghost Membrane Preparation

Fresh red blood samples were centrifuged at 15000xg for 10 mins to get rid of plasma and the buffy coat and precipitate the red blood cells (RBCs). After that, 5 mL of the RBCs were incubated with 3mg/ml as a final concentration of bromelain at 37°C for 1 hr to cleave the 6th external loop of RhD protein on the RBC membranes.

Then, samples were washed three times with 0.15 M NaCl solution. 0.2 M Monobasic and 0.2 M Dibasic solution were prepared by dissolving NaH₂PO₄.2H₂O and anhydrous Na₂HPO₄ respectively in 1 L of distilled water. To prepare the stock solution of phosphate buffer (0.1M Na₂PO₄ pH8.0) monobasic and dibasic solutions were mixed together.

The samples were washed with 0.1M Na₂PO₄ pH8.0 until the supernatant was clear followed by incubating the pellet for 10 mins at 4°C. Samples were transferred into centrifuge tubes. Samples were mixed thoroughly and were incubated for 5 mins at 4°C with 5M 0.1M Na₂PO₄. After that, these samples were centrifuged at 16000xg at 4°C for 5 mins, the supernatant was removed. This step was repeated until the membranes were creamy pink. Membranes were transferred to 1mL Eppendorf tubes.

1 mL of 5 mM phosphate buffer was added and membranes were centrifuged at 22 xg for 10 mins. Clear supernatant was discarded followed by adding 100 µL 5mM phosphate along with 10

μ L 2mM phenylmethanesulphonyl fluoride (Sigma-Aldrich, UK) to the membrane to prevent proteolysis of samples during long term storage. Samples were stored at -80°C until required.

2.2.3.15 Sodium Dodecyl Sulphate-polyacrylamide Gel Electrophoresis (SDS-PAGE)

5 μ g of the protein per sample from bacterial protein and 10 μ g from the human ghost membrane samples were used. To prepare the samples in order to load them into a gel, samples were mixed with 4x loading sample buffer (50% (w/v) glycerol, 5% (w/v) lithium dodecyl sulphate, 5% (w/v) Tris(hydroxymethyl) aminomethane, 1.0% (v/v) hydrochloric acid) (BioRAD, U.K) by a ratio of 3:1, sample: loading dye.

Then, 2 μ L of XT reducing agent (BIO-RAD, UK) were added (BioRAD, U.K) to each sample. Next, the samples were heated in a hot block at 95°C for 5 mins prior loading them onto the gel. Samples were loaded into pre-cast gel XT Bis-Tris 4-12%, 12+2 well, 45 μ L , Criterion Cell (BioRad, U.K). BlueEye Marker (GeneFlow, U.K) and Precision Plus ProteinTM WesternCTM Standards were used as a marker to determine the protein molecular weight. Then, 3-(N-morpholino)propanesulfonic acid running buffer 1x XTMOPS (BIO-RAD, UK) was prepared (1:20). Gels were run at 150V for 90 mins or until the dye went to the lowest point of the gels. Gels were either stained or immunoblot-

ted (Western blotted).

2.2.3.15.1 Instant Blue Staining

To test the quality control of the gel electrophoresis and to check the presence of the proteins in samples, the gel was immersed overnight at room temperature on a shaker in Instant Blue stain (Expedeon, U.K). Next day, the gel was washed and was imaged by a phone camera.

2.2.3.15.2 Western Blotting

Polyvinylidene fluoride (PVDF) membrane with 0.45 μm pore size (Merck Millipore, U.K) was activated by submersing the membrane in Methanol for 20 sec followed by 2 min in dH_2O then in cold transfer buffer (190 mM glycine, 25 mM Tris base, and 20% (v/v) Methanol).

To prepare the sandwich of the membrane, layers of the sponge, filter paper, gel peice (which contain proteins and the marker), the prepared PVDF membrane, second filter paper, and finally the second sponge layer then this was placed into the cassette of the wet transfer blotting Bio-Rad system for 60 minutes at 100 V. After the proteins were transferred from the gel to the PVDF membrane, this membrane was placed in sterilized plastic tray with lid, blocking solution (1x PBS-Tween 20 and 5% (w/v) of dried milk powder) was used to block the membrane for an hour at room temperature on the shaker. After that, the blocking buffer was discarded and the primary antibodies were added either neat or diluted in the blocking buffer with (1:6) ratio, to the membrane as the manufacturer's instructions recommended. Then, it was stored on a shaker in the cold room for overnight. On the next day, three washes with 1x PBS-T were to the PVDF membrane each wash was 10 min long at room temperature on a shaker. Secondary antibody was prepared with a ratio of 1:1000 according to manufacturer's instructions. It was added to the membrane after discarding the 1x PBS-T solution. Next, PVDF

membrane was incubated for 1 hour at RT followed by washing three times with 1x PBS-T and then, one wash with 1x PBS; each wash was 10 min long at room temperature on a shaker. Finally, PierceTM ECL Plus Western Blotting Substrate (Thermo Fisher, U.K) was prepared according to manufacturer's instructions and added to the PVDF membrane before taking the images of the membrane by using PXi6 Imaging System (Syngene) to enhance the membrane chemiluminescent visualization.

However, 10 μ g of protein from ghost membrane with different Rh phenotypes and 5 μ g of protein from all constructs and wild types samples were run using 1D gel electrophoresis prior to transfer them onto PVDF membrane. Moreover, cultural supernatant 2.2B9 were used at the concentration 1:6 as a primary antibody followed by rabbit anti-mouse horse radish peroxidase as a secondary antibody.

2.2.3.16 Flow cytometry

2.2.3.16.1 Flow cytometry for Blood Samples

RhD⁺, RhD⁻ blood samples were ordered from National Health Service (NHSBT) and were used as a control to test the efficiency of our Human antibodies; 2.2B9 which reacts against human RhD external loop 4, BIRMA-D6 which reacts against human RhD external loops 4, and 6, and MS-26, which reacts against human RhD external loops 3, 4, and 6.

Whole blood was washed in 30 ml of 1x PBS and it was centrifuged for 5 min at 15000 xg then the supernatant were discarded, this step was repeated three times. 20 μ L of the final pellet from the blood samples was used and were diluted in 30 ml of 1x PBS. Then, 200 μ L of this mixture was used to mix it with 100 μ L of tissue culture supernatant human antibodies (2.2B9 and MS-26) and 10 μ L of the purified human antibody(BIRMA-D6). These samples were incubated for 1 hour at 4 °C in dark. PBS were added followed by centrifuge samples for 5 min at 15000 xg at room temperature, this step was repeated for two times. Then, 200 μ L of 1x PBS was used to resuspended the pellet. 10 μ L of secondary FITC antibody was added and incubated with samples for 30 min at 4 °C in dark. Samples were washed with 1x PBS and centrifuged at 4 °C for two times at 15000 xg. 300 μ L of 1x PBS was used to resuspend the samples. Finally, samples were transferred into FACS tubes and analysed by FACS ARIA II Flow cytometer/Cell Sorter (BD, U.K)

2.2.3.16.2 Flow cytometry for *E. coli* Broth

5 ml of LB broth with 50 μ L Kn were grown overnight at 37 °C at 200 rpm for the positive colonies from the last transformation of PET vector/ *NeRh50* with human external loop (6/ 4,6/ 3,4,5/ 2,3,4,6). 5 ml LB broth with Kn for BL21 ED3 competent cells, 5 ml LB broth with Kn for *Nitrosomonas europaea* ATCC 19718 were used as a positive control.

All broth were harvested on 22000 xg at room temperature for 10 min, then, supernatant were discarded. Pellets were washed with 1 ml of cold 1x PBS. After that, samples were centrifuged at the maximum speed at room temperature for 10 min, the supernatant were discarded. This step was repeated for two times. 700 μ L of resuspending mixture [5% (w/v) Triton X-100, 0.15 (v/v) NaCl, 10mM Tris-HCl, and 1mM EDTA pH 8.0] were used to resuspend pellets.

50 μ L of 10 mg/ml of freshly prepared Lysozyme were added to the previous mixture and vortexed for sec. The final mixture were incubated for cell wall lysis step for 30 min at 37 °C. Followed with spinning at 1500 xg at room temperature for 10 min, supernatant were discarded. Then, all pellets were resuspended with 1 ml of cold 1x PBS. This mixture were divided in to 4 main tubes; unstain sample, only secondary, only secondary with blocking buffer and primary Abs.

100 μ L of culture supernatant Abs, 10 μ L of purified Abs; Anti-D (2.2B9, BIRMA-D6, MS-26)/ Anti-NeRh50 (1918, 1755, 1917) and Anti-*E. coli* were added to the right tube. Then, these tubes were incubated in dark at 4°C for 45 min. Samples were centrifuged at the maximum speed at room temperature for 10 min, the supernatant were discarded. Pellets were washed with 1 ml of cold 1x PBS. This step was repeated two times.

500 μ L of 1% BSA (blocking buffer) were added to the right tubes while 500 μ L of cold 1x PBS were added to the other tubes. Then, these tubes were incubated in the dark on shaker at 4°C for 1hr. Samples were centrifuged at 22000 xg at room temperature for 10 min, supernatant were discarded. Pellets were washed with 1 ml of cold 1xPBS. This step was repeated two times.

10 μ L of the suitable secondary- FITC conjugated Ab (Mouse anti human FITC, Goat anti mouse FITC, and Goat anti rabbit FITC) were added and incubated in dark at 4°C for 1hr then, washing step was repeated twice.

Finally, 1 ml of cold 1x PBS were used to resuspend pellets to prepare the sample to be read by BD FACS Aria II - Flow cytometer/Cell Sorter which was purchased from Beckman Coulter life sciences, UK.

Chapter 3

Results

3.1 Introduction

Blood groups were discovered in 1901 and there are plenty of studies in this area, reflecting the importance of this science in the Immuno-haematology field. The antigen/antibody reaction is one of the most significant features in serology. Due to the importance of this approach, all hospitals and blood banks worldwide rely on the phenotyping of blood groups as one of the routine blood tests.

The disadvantage of the current antibody screening and identification methods is the use of human red blood cells, as these can complicate the detection of antibody in different cases for example, if more than one antibody specificity is present, in the case of multi-transfused individuals and in the case of having weak antigens. Moreover, using human red blood cells, that have a limited shelf life, is another barrier to study rare blood phenotype cases.

Thus, to overcome the downsides of using human red cells, in this project, the aim was to construct recombinant Rh antigens. The concept was that if bacteria can express Rh antigens, they may reduce the need for human red cells in complicated cases. The principle of using blood group recombinant antigens to identify blood group antibodies in human serum is well affirmed, and has been achieved for many blood group systems but not for Rh blood group system (Ridgwell *et al.*, 2007).

The Rh blood group system is the second most important blood group system in transfusion medicine (Westhoff, 2004). Anti-Rh system antibodies are major causes of transfusion reactions and haemolytic disease of the fetus and newborn (Van Kim *et al.*, 2006). Despite the great importance of Rh blood group system, there has not been any successful production of recombinant Rh antigens in prokaryotic cells (Conroy *et al.*, 2005, (Ridgwell *et al.*, 2007).

Rh blood group system has five genes, the blood group active *RHCE* and *RHD*, the erythroid ammonium transporter *RHAG* and two non-erythroid counterparts, *RHBG* and *RHCG*. Rh-related proteins are expressed in all animals, humans, bacteria (including AmtB) having a highly conserved membrane domain, termed the Ammonium transporter fold (Marini *et al.*, 2000). The Rh recombinant antigens would produce a fast, lower bio-hazard, affordable technique depending on the expression of RhD protein in the bacterial membrane.

The recombinant gene is composed of *NeRh50* backbone and multiple combinations of human RhD external loops. Firstly, we attempted proof of concept by designing a *NeRh50"6* construct, with the location of loop 6 being between two highly conserved transmembrane domains (11-12), which increases the probability of this gene sequence being expressed on the *E. coli* inner membrane.

Choosing to work with the external loops (6, 4, and 3) to cre-

ate the recombinant gene constructs was based on the successful work of (Liu *et al.*, 1999 a, 1999 b). These studies achieved Rh expression by constructing the recombinant *RHCE RHD* cDNAs by replacing the human *RHCE* external loops 3, 4, and 6 with the corresponding human *RHD* external loops 3, 4, and 6 sequences to construct a cDNA expressing a hybrid form of Rh protein, which has a backbone of RhCE and external loops from human RhD. Besides these results, in 1997, there was another study by Avent *et al.*, where they were able to determine the 9 critical amino acids (AA) for expression of the D epitopes, which are located on the predicted external loops 3, 4, and 6 of the RhD protein. These critical amino acid are: Met169, Met170, Ile172, Phe223, Ala226, Glu233, Asp350, Gly353, and Ala354. Furthermore, in 2005, Conroy *et al.*, have successfully constructed homology models of RhD and RhAG based on using the predicted secondary structure of Rh proteins (Figure 1.8) and the structure of *E. coli* AmtB (Figure 1.9) as a template. These models predicted the human RhD external loops, which assessed constructing the hybrid *NeRh50*” human RhD external loops (2, 3, 4, and 6). The entire Rh complex comprising one Rh protein monomer and two RhAG subunits form a trimer arranged in a cloverleaf configuration (Conroy *et al.*, 2005; Callebaut *et al.*, 2006).

Choosing the *N. europaea Rh50* like protein to be the backbone of the constructed gene instead of another bacterial species such

as *E. coli* AmtB is due to the similar topology of the *NeRh50* like protein to human Rh proteins. *NeRh50* has 6 external domains with 12 membrane domains (Figure 1.13) like human Rh proteins, while *E. coli* AmtB protein has 5 external loops with 11 membrane domains (Figure 1.11). The similarity in the topology between human RhD and *N. europaea* Rh50 like protein is more likely to allow the human RhD epitopes to be expressed. Moreover, *N. europaea* was also chosen because it was a prokaryote, before this study no successful expression of eukaryote Rh cDNAs have been described in prokaryotic cells.

3.1.1 Aim of the study

Aims of this chapter will be divided in two main section based on the results of this project,

A) Aims for generating constructs:

- Map the human RhD external loops (2, 3, 4, and 6) with the corresponding domains in the *NeRh50* protein.
- Generate a recombinant gene composed of *NeRh50* backbone and human RhD external loops (2, 3, 4, and 6) to create; *NeRh50*"6, *NeRh50*"4, 6, *NeRh50*"3, 4, 6, and *NeRh50*"2, 3, 4, 6.
- Clone these constructs in to a suitable expression vector.
- Produce these constructs in a prokaryotic expression system, which is *E. coli*.

- Confirm the presence of all constructs in the pET vector by enzymatic digestion.

B) Aims for confirming the presence of constructs and their efficiency by testing for proteins expression:

- Confirm the presence of the hybrid recombinant genes at the protein level by detecting their protein expression on the inner *E. coli* membrane.
- Use Several polyclonal and monoclonal antibodies (Table 2.1).
- Use different techniques; Western blot, Flow-cytometry, immunogold staining and Transmission electron microscopy (TEM) (Table 3.2) to assess the presence of the *NeRh50* human RhD external loops in the *E. coli* membrane.

So for that, the aim of this project is to generate Rh recombinant antigens by constructing hybrid recombinant *N. europaea-human RH* genes in order to prove that the recombinant RhD proteins can mimic the Rh blood group antigen (D) as a novel approach.

3.2 Constructing *RH* recombinant genes for expression in *E. coli* membrane.

3.2.1 Mapping human RhD 6th external loop

Mapping and alignment of the various 13 Rh homologous proteins shows a high degree of variation between the 6th external loop sequence (Figure 3.1).

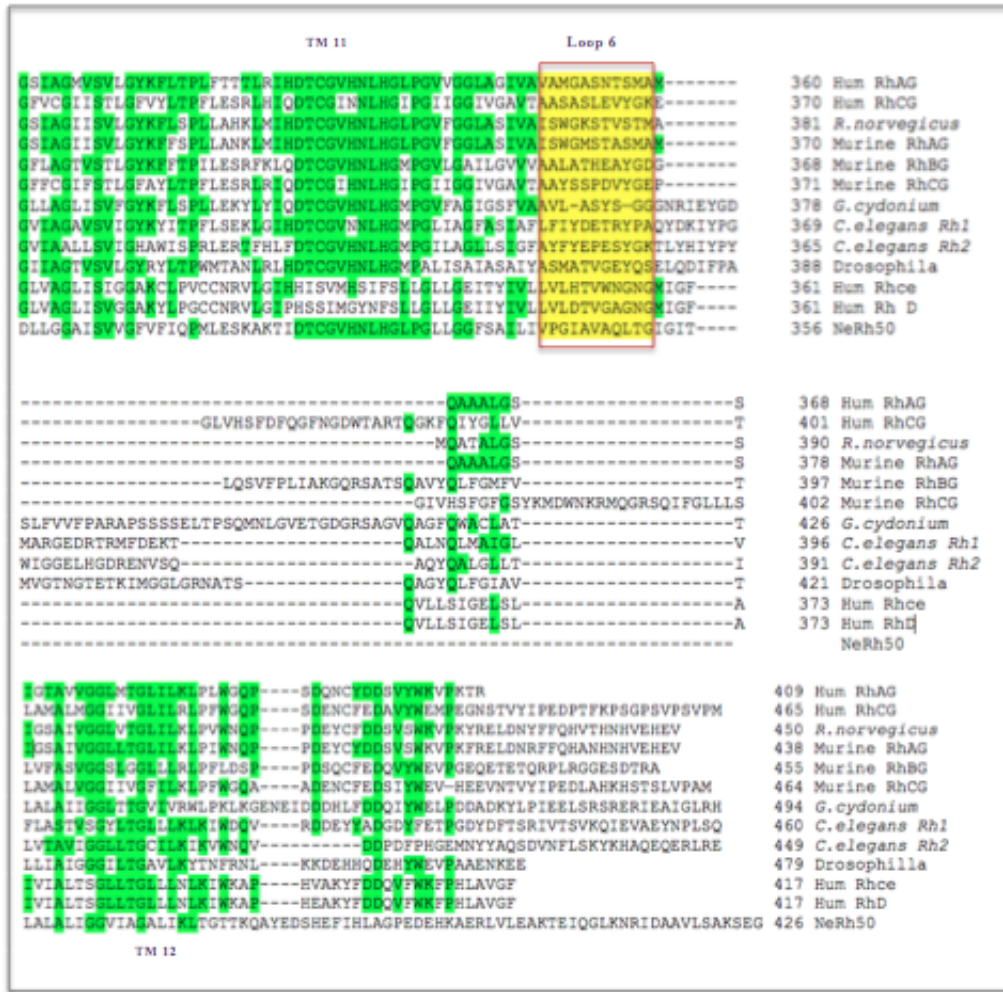


Figure 3.1. *Clustal X* alignment of the Transmembrane Domain (TM) (11-12) region of 13 Rh proteins from different species.

The yellow highlighted sequence shows the external loop 6, while the green highlighted AA residues are showing the consensus AA in the Rh protein family homology. This alignment shows that the transmembrane domain (11-12) is highly conserved between the Rh protein family in different species, while loop 6 has different AA sequences between these species.

This alignment shows a massive variation in loop 6 structure between different species, but the 6th loop, which is located between two highly conserved transmembrane domains; 11 and 12. Therefore, this doesn't necessarily follow but as the TM domains are conserved, we had hoped the replacement of the loop part would not disrupt expression of hybrid Rh protein from NeRh50

and human RhD loop 6 in *E. coli*.

GenScript Codon Usage Frequency Table (chart) Tool. [online] Available at: <http://www.genscript.com/tools/codon-frequency-table> [23-06-2019] was used to reverse translate the AA sequence of human RhD external loops; loop 6 [LVLDTVGAGNGM] (Figure 3.2), loop 2 [FPSGKVVIT] (Figure 3.7), loop 3 [NIFNTDYH MNMM] (Figure 3.8) and loop 4 [ALLRSPIERK] (Figure 3.9), back to the most likely nucleotide sequence.

Therefore, minimizing the mutation incidence in the bacteria by using the highest frequency codons, which will make the recombinant protein sequences more likely to be successfully expressed by the *E. coli* (Table 3.1).

The AA sequence of external loop 6 of the human RhD (Figure 3.2) was used to replace the corresponding loop 6 sequence in *NeRh50* (Figure 1.13) to generate the predicted *NeRh50"6* topology (Figure 3.3).

The loop 6 amino acid sequence was reverse translated to nucleotides by using the *E. coli* codon usage table website, GenScript. (2002-2019). GenScript Codon Usage Frequency Table(chart) Tool. [online] Available at: <http://www.genscript.com/tools/codon-frequency-table> [23-06-2019].

These nucleotides were added to the *NeRh50* backbone. The nucleotide sequence was sent to GenScript Company, USA to synthesis the cDNA of the first construct: hybrid recombinant gene of *N. europaea Rh50* with a human *RhD* external loop 6 (*NeRh50"6*).

1	2	3	4	5	6	7	8	9	10	11	12
L	V	L	D	T	V	G	A	G	N	G	M
CTG	GTG	CTG	GAT	ACC	GTT	GGT	GCC	GGC	AAC	GGT	ATG

Figure 3.2. Amino acid sequence of Human RhD external loop 6 and the corresponding nucleotide sequence based on *E. coli* codon usage.

Red numbers on the above show the order of this AA in the 6th loop of human RhD protein sequence. Bottom box shows the reverse translation of this AA sequence back to the nucleotide level based on the *E. coli* codon usage table.

Amino Acid	3-letters Abbreviation	1-Letter Abbreviation	Codon	Frequency
Leucine	Leu	L	CTG	52.9
			CTA	4.2
			CTC	10.2
			CTT	11.9
			TTC	13.0
			TTA	14.3
Aspartic Acid	Asp	D	GAT	38.6
			GAC	19.2
Threonine	Thr	T	ACA	22.8
			ACG	13.7
			ACC	20.0
			ACT	10.3
Valine	Val	V	GTT	19.9
			GTC	14.3
			GTA	11.6
			GTG	24.4
Glycine	Gly	G	GGT	22.2
			GGA	9.5
			GGC	27.1
			GGG	11.3
Alanine	Ala	A	GCA	21.8
			GCG	30.1
			GCC	24.2
			GCT	17.1
Asparagine	Asn	N	AAC	21.2
			AAT	20.6
Methionine	Met	M	ATG	23.9
Tyrosine	Tyr	Y	TAT	17.5
			TAC	12.2
Phenylalanine	Phe	F	TTT	22.1
			TTC	16.0
Arginine	Arg	R	AGA	3.6
			AGG	2.1
			CGT	20.0
			CGC	19.7
			CGA	3.8
			CGG	5.9
Serine	Ser	S	AGT	9.9
			AGC	13.2
			TCT	10.4
			TCC	9.1
			TCA	8.9
			TCG	8.5
Proline	Pro	P	CCT	7.5
			CCC	5.4
			CCA	8.6
			CCG	20.9
Isoleucine	Ile	I	ATT	29.8
			ATC	23.7
			ATA	6.8
Glutamic acid	Glu	E	GAA	39.1
			GAG	18.7
Lysine	Lys	K	AAA	35.3
			AAG	12.4
Histidine	His	H	CAT	12.5
			CAC	9.3

Table 3.1. List of human RhD external loops AA and *E. coli* nucleotide codons of these AA.

This table shows codon of each AA in human RhD external loops (2, 3, 4, and 6) based on the codon usage table of *E. coli* website, GenScript. (2002-2019). GenScript Codon Usage Frequency Table(chart) Tool. [online] Available at: <http://www.genscript.com/tools/codon-frequency-table> [23-06-2019].

The yellow highlighted codons show the highest frequency and those were used to reverse translate the AA.

Figure 3.3. Predicted membrane topology and amino acid sequence of *NeRh50"*6 construct.

This figure depicts the hypothetical membrane orientation of the recombinant *NeRh50"6* in which predicted externalized amino acids are numbered. The recombinant protein was derived initially from a cDNA of a wild type NeRh50 combined with a human RhD external loop 6 polypeptide sequence. The predicted extracellular loop of the human RhD amino acids [LVLDTVGAGNGM] are shown as solid circles with a blue single letter amino acid code inside each circle. The yellow arrow point to the external loop 6.

NeRh50''6 hybrid recombinant gene size was 1766 bp and it was inserted in the pUC57 vector (Figure 3.4). Thereafter, pUC57 plasmid-*NeRh50''6* was transformed into *DH5 α E. coli* competent cells.

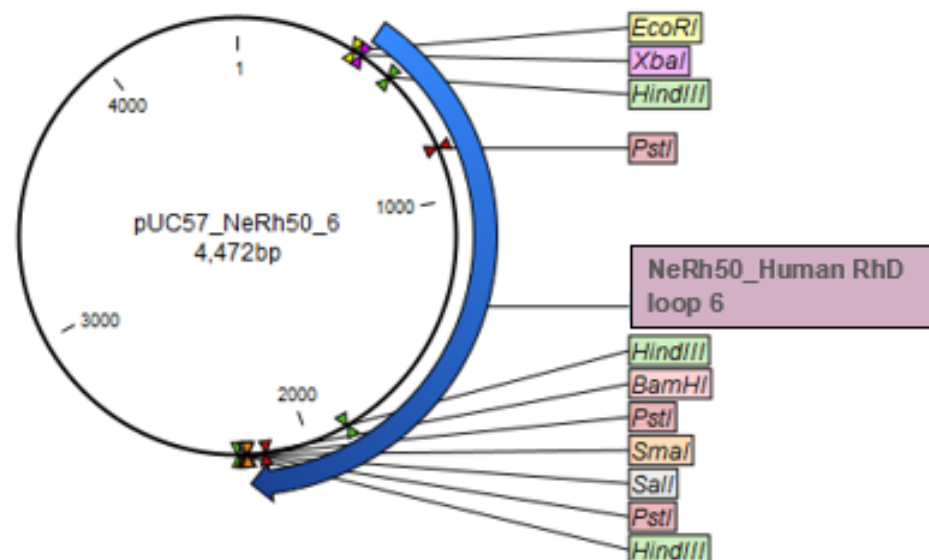


Figure 3.4. pUC57 vector with the recombinant gene (*NeRh50"6*) circular sequence map.

Is a circular sequence of the full pUC57 vector with the recombinant gene (*NeRh50"6*), it shows all the restriction enzymes sites. By CLC Genomics Workbench 12.0 programme.

The sequence in Figure 3.5 was used to map the AD1/ AD2 primers and *Bam*HI and *Xba*I restriction sites in the pUC57-recombinant gene (*NeRh50''6*) plasmid sequence. *Bam*HI and *Xba*I restriction enzymes were used to digest the hybrid recombinant gene *NeRh50''6* from the pUC57 vector. The digestion plasmid shows 2 bands (1.7 and 3 kb). The lower band (1.7 kb) is the correct size for our insertion (1766 bp) (Figure 3.6).

1 I R A R Y L A N A S R * S C K W K V K C
 1 ATTCGAGCTCGGTaCCTGCGGAATGCA**TCTAC**TAGAGTTGTAAATGGAAAGTGAATGT
 1 10 20 30 40 50
 1 TAAGCTCGAGCCATGGAGCGCTTACGTAGATCTATCTCAACATTTAOCCTTCCACTTTACA

 21 * F L L G S L * Q N S Q T I I Y L * A P
 61 TAATTCCTGTAGGCGAGCTCTGACAGAACAGCCAGACATCATTATTATGAGCGGCC
 61 70 80 90 100 110
 61 ATTAAAGGACAATCCGTGGGAGACTGTCTTGTGGTCTGTTAGTAATAAATACTCGCGGG

 41 K L R R T W * I T I E I R I P Q * N R Y
 121 AAGCTTCGTGGGACCTGCTAAATAACAATAGAAATCAGAAATACCGCAGTGAATAGATAT
 121 130 140 150 160 170
 121 TTCGAAGCAGCCTGGAGCCATTTATTGTTATCTTTAGTCTTATGGCGTCACCTTTATCTATA

 61 G I R F S L I Y A E N I P V * * T D Q M
 181 GGAATCAGGTTTTCGCTTATTATGCGGAAATATTCCTGTTAATGAACTGATCAAAATG
 181 190 200 210 220 230
 181 CCTTAGTCCAAAAGCGAATAAATACGCCCTTTTATAAGGCAAAATTACTTGAAGTACTTTAC

 81 R L I L A P * L Y Q P E S C R H I R G V
 241 AGGTTGATCTTAGCACCATAGCTGTATCAACAGAAATCCGGCAGGCACATACGGCGGGTA
 241 250 260 270 280 290
 241 TCCAACTAGAAATCGTGGTATCGACATAGTTGGTCTTAGGCCGCTCCGTGTATGCGCCCAT

 101 T R I T S I * D I G S Q Q M S K H L C F
 301 ACCCGGATAACTTCAATTTAGATATA**GGCACTCAACAAAGCACTAAACGCTATCTTTG**
 301 310 320 330 340 350
 301 TGGGCTTATGGAAGTTAAATCTATATCCCTCAGTTGTTTACTCATTTCGTGGATACAAAG

 121 T A F S S I A L F L L C F S S W A S A V
 361 **ACTGCGTTCAGTTCAATCGCTCTGTTTCTGTTGTTGTTTCAGCACTTCGGCGAGCGCACT**
 361 370 380 390 400 410
 361 TGACGCAAGTCAAGTTAGCGAGACAAAGACAAACAAAGTCTCAACCGGCTCGCGTCRA

 141 A P A E I N E A R L V A Q Y N Y S I N I
 421 **CGTCTGACAGATCAACCAAGCGCGCTCTGGTAGCGCAATATAATTACTCCATAAACAT**
 421 430 440 450 460 470
 421 CGAGGACGCTCTCTAGTTGCTTCGGGACAGACCTCGCGTTATATTAAATGAGGTATTTGTAA

 161 L A M L L V G F G F L M V F V R R Y G F
 481 **CTGCGCATGCTGCTGCTCGCTTTCGCTTTTAAATGCTTTTCCTACCAAGCTACCGGTT**
 481 490 500 510 520 530
 481 GACCGGTACGACGACCGACGCAAGGCAAAATTAACCAAAAGCATGCTTCCATGCCCAA

 181 S A T T G T Y L V V A T G L P L Y I L L
 541 **AGCGCAACTACCGCTACCTATCTGCTGCTAGCCACCGCACTTCCACTCTATATTCTTTT**
 541 550 560 570 580 590
 541 TCGCGTTGATGGCCATGGATAGACCGCATCGGTGGGCTGAAGGTGAGATATAGAAAT

 201 R A N G I F G H A L T P H S V D A V I Y
 601 **CGCGCAACCGGATTTTCGGGACCGCATTGACGCGCCACTCCGTTGATGCACTAATCTAT**
 601 610 620 630 640 650
 601 GCGCGTTTGCCTAAAGGCCCGTGGCTAACTGCGGGGTGAGGCACTACGTCTATTAGATA

 221 A E F A V A T G L I A M G A V L G R L R
 661 **CGCAATTTGCGGTGCTACGGGCTGATTGCGCATGCGCACTACTCGGGGCTTTACGG**
 661 670 680 690 700 710
 661 CGGCTTAAAGCGCACCGATGCCCGACTAACGGTACCGCTGCTCATGAGCGCGCAATGCC

 241 V F Q Y A L L A L F I V P V Y L L N E W
 721 **CTTTTCAATACCGCTTACTGCCATTCTTTATCGTGGCGGTTTATCTGCTCAATGAGTGG**
 721 730 740 750 760 770
 721 CAAAAGTTATGCGCAATGACCGTAAACAAATAGCACGGCCAAATAGACGACTTACTCACC

pUCS7

261 L V L D N A S G L T E G F Q D S A G S I
 781 CTGCTGCTGGTAATGCTTCAGGGCTTCACTGAAAGGATTCAGGATTCGCGCGATCGATC
 781 790 800 810 820 830
 781 GACCACGACCTATTACGAAGTCCCAACTGACTTCCTAAGGTCCCTAAGACGGGCTAGCTAG

 281 A I H A F C A Y F C L G V S I A L T T A
 841 CCGATCCATGCTTTTGGCTGCATATTTTGGTTTGGAGTCTCCATCGCTCTGACTACTGCG
 841 850 860 870 880 890
 841 CGATAGGTACGAAAGACCGTATAAACCAACCCCTCAGAGGTAGCGAGACTGATGACGC

 301 A Q R A Q P I E S D A T S D R F S M L G
 901 GCGCAGCGGGCTCAGCGGATCGAATCGGATCGGACTTCGGATCGCTTTTCAATGCTGGG
 901 910 920 930 940 950
 901 CGGCTCGCGGAGTGGCTAGCTTAGGCTACGCTGAAGGCTAGCGAAGATTACGACCT

 321 S M V L W L F W P S F A T A I V P F E Q
 961 TCAATGCTACTGTGGCTTTTCTGGCGGAGCTTTGGGACTGCAATGCTACCATTCGAAACG
 961 970 980 990 1000 1010
 961 AGTTACCATGACACCGAAAGACCGGCTCGAAACGCTGACGTTAGCATGCTAAGCTTGTG

 341 M P Q T I V N T L L A L C G A T L A T Y
 1021 ATGCCCCAAACCATAGTCAACACACTGCTTGGCTTATGCGGGCGGCACACTGCGCACTAT
 1021 1030 1040 1050 1060 1070
 1021 TACGGGCTTTGGTATCAGTTGTGTGACGAAACGAAATACGCGCGGCTGTGACGCGGTGATA

 361 F L S A L F H K G K A S I V D M A N A A
 1081 TTGCTGAGCGCATGCTTTTCAAAAGGAAAGCATCCATAGTACCATGCGGCAATCGACG
 1081 1090 1100 1110 1120 1130
 1081 AAGGACTCGCGTACAAAGTGTTTCCTTTTCGTAGGTATCATCTGTACCGGTTACGTCGT

 381 L A G G V A I G S V C N I V G P V G A F
 1141 CTTGCTCGCGGGCTAGCGATCGGGTCACTCTGTAAATGCTGGGACCACTAGGTGCTTTT
 1141 1150 1160 1170 1180 1190
 1141 GAACGACCCCCCATCGCTAGCGCGAGTCAGACATTTGTAGCAGCTCGGTATCCACGAAA

 401 V I G L L G G A I S V V G F V F I Q P M
 1201 CTGATTTGGTTTGGCTGGGTGCTATTCTGCTGGGCTTTGGTATTTATTCAGCCCATG
 1201 1210 1220 1230 1240 1250
 1201 CACTAACCAGACGCCACCCACGATAAAGACAGCCACCCAGACATAAATAAGTCGGGTAC

 421 L E S K A K T I D T C G V H N L H G L P
 1261 CTGGAATCCAAAGCCAGACGATAGATACCTGGCGGGTACACAACTCTGCATGGATTACCA
 1261 1270 1280 1290 1300 1310
 1261 GACCTTAGGTTTTCGGTCTGCTATCTATGAGCGGCCCATGCTGTAGAGGTACCTAATGCT

 441 G L L G G F S A I L I L V L D T V G A G
 1321 CGATTATTGGCGGATTTAGTGGATGCTGATACTGGCTGCTGGATAACCGTGGCGCGCGCG
 1321 1330 1340 1350 1360 1370
 1321 CCTAATAACCGGCTAATCAGCGCTAGGACTATGACCAAGACCTATGGCACCGCGCGCGCG

 461 N G M G I T L A L A L I G G V I A G A L
 1381 AACCCCATCGGTATCACACTTGGCGTGGCTTTGATTGGCGGTGTGATTGCTGGAGCACTG
 1381 1390 1400 1410 1420 1430
 1381 TTGCGGTACCCATAGTGTGAACGCGACCGAACTAACCGCCACACTAACGACCTCGTGAC

 481 I K L T G T T K Q A Y E D S H E F I H L
 1441 ATCAAACTAACCCTTACACCGAAACAGCTTATGAGGACAGTCACGAGTTTATCCATCTT
 1441 1450 1460 1470 1480 1490
 1441 TAGTTTGATTGGCCATGTTGCTTTGTTTCGATACCTGCTGCTGCTCAATAGGTAGAA

 501 A G P E D E H K A E R L V L E A K T E I
 1501 GCAGGACCTGAGATGAGCATAAAGCTGAAGGCTGGTATTAGAGGCAAAACTGAAATA
 1501 1510 1520 1530 1540 1550
 1501 CGTCCTGGACTTCTACTGCTATTTGACTTTCCGACCATATCTCCGTTTTTGACTTTAT

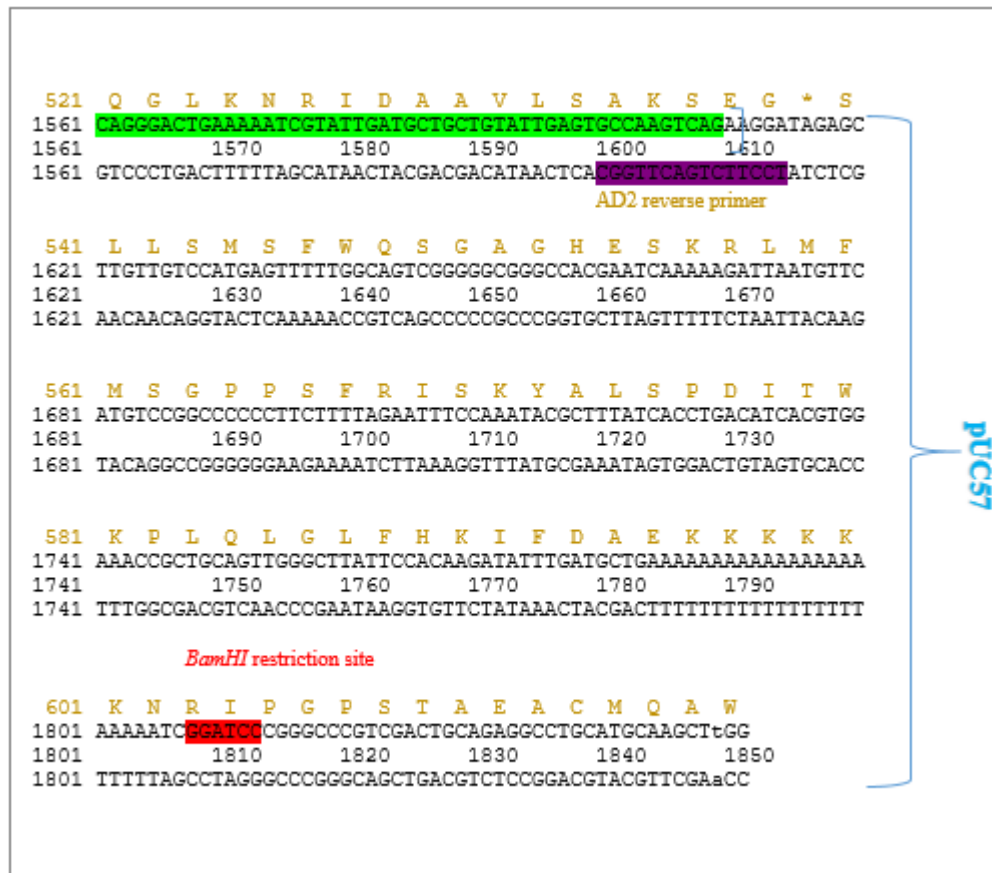


Figure 3.5. pUC57 linear vector sequence with the recombinant gene (*NeRh5026*).

A linear nucleotide sequence of part of the pUC57 vector with the recombinant gene (*NeRh5026*). Blue highlighted nucleotides sequence refers to *XbaI* restriction site. Green highlighted nucleotides sequence refers to *Nitrosomonas europaea* Rh50. Red nucleotides sequence refers to human RhD loop 6. Red highlighted nucleotides sequence shows *BamHI* restriction site. Yellow highlighted nucleotides sequence is *AD1* forward primer. Purple highlighted nucleotides sequence is an *AD2* reverse primer.

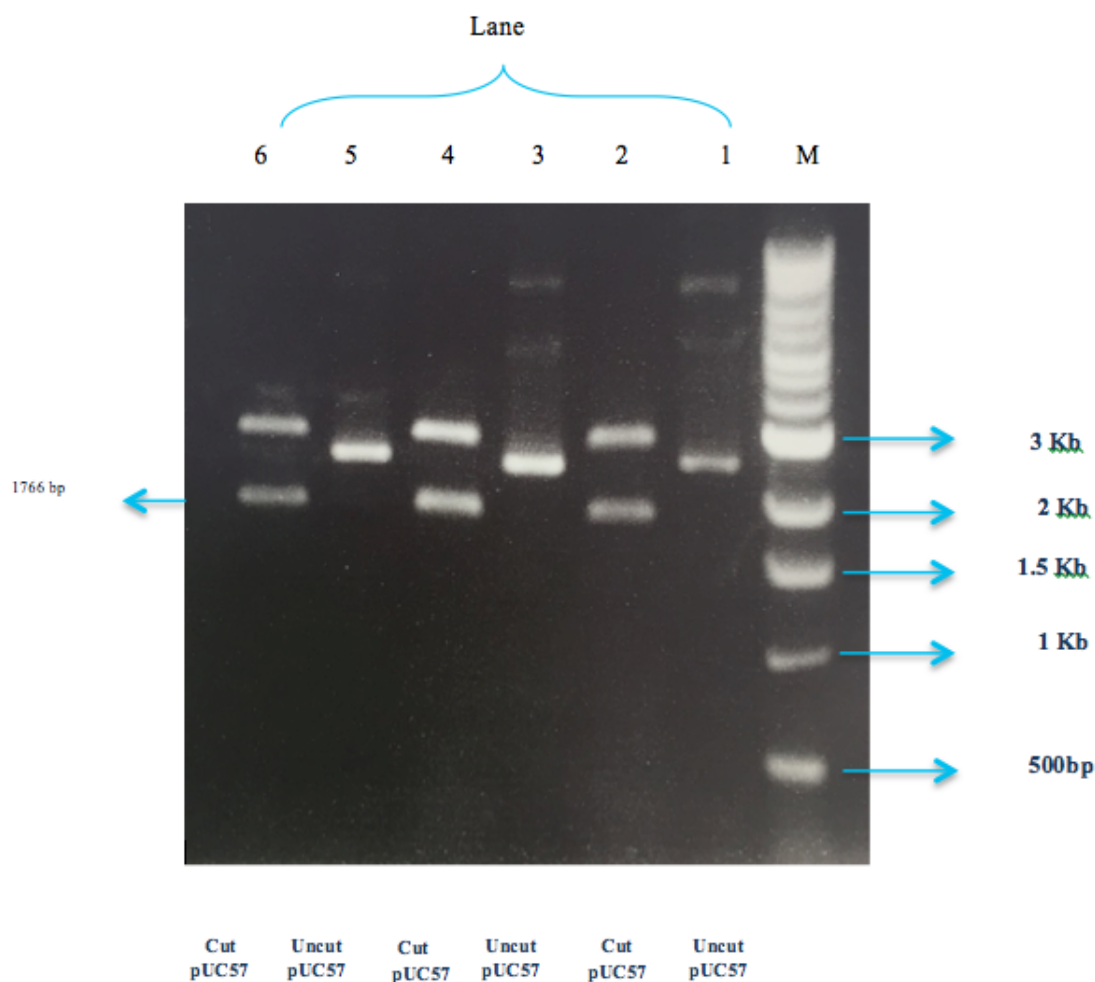


Figure 3.6. *Bam*HI and *Xba*I digestion of pUC57-*NeRh50''6*.

Lane numbers (1, 3, and 5) are representing three colonies of transformed pUC57 plasmid with the recombinant gene *NeRh50''6*, (non-cut sample), which is the samples not treated with the *Bam*HI and *Xba*I restriction enzymes. Lane numbers (2, 4, and 6) are representing the digested pUC57 plasmid (the cut sample) with restriction enzymes (*Bam*HI and *Xba*I) cut the plasmid. Band sizes are 3 kb and 1.7 kb. The lower band (1.7 kb) represents the constructed recombinant gene. Products were run on a 1% (w/v) Agarose gel alongside a DNA marker (M). Shown to the right of the figure are marker sizes in base pairs.

After testing *NeRh50*"6, same steps were followed to replace other external loops of the *NeRh50* wild type with human RhD corresponding external loops; human RhD loop 2 (Figure 3.7), human RhD external loop 3 (Figure 3.8), and human RhD external loop 4 (Figure 3.9) were used to replace the corresponding loop in wild type *NeRh50* (Figure 1.13) to generate the other constructs of *NeRh50* wild type backbone with different human RhD external loops topology; *NeRh50*"4, 6 (Figure 3.10), *NeRh50*"3, 4, 6 (Figure 3.11), *NeRh50*"2, 3, 4, 6 (Figure 3.12). *NeRh50*"4, 6, *NeRh50*"3, 4, 6, and *NeRh50*"2, 3, 4, 6 hybrid recombinant genes size were 1275 bp and they arrived inserted in the pUC57 vector from GenScript (Figure 3.13).

Thereafter, all the constructs [pUC57 plasmid-*NeRh50*"4, 6 / pUC57 plasmid-*NeRh50*"3, 4, 6 / pUC57 plasmid-*NeRh50*"2, 3, 4, 6] were transformed into *DH5 α* E. coli competent cells.

1	2	3	4	5	6	7	8	9
F	P	S	G	K	V	V	I	T
TTT	CCG	AGC	GGC	AAA	GTG	GTT	ATC	ACC

Figure 3.7. Amino acid sequence of Human RhD external loop 2 and the corresponding nucleotide sequence based on *E. coli* codon usage.

Red numbers on the above show the order of this AA in the 2nd loop of human RhD protein sequence. Bottom box shows the reverse translation of this AA sequence back to the nucleotide level based on the *E. coli* codon usage table.

1	2	3	4	5	6	7	8	9	10	11	12
N	I	F	N	T	D	Y	H	M	N	M	M
AAC	ATT	TTC	AAC	ACC	GAC	TAC	CAC	ATG	AAC	ATG	ATG

Figure 3.8. Amino acid sequence of Human RhD external loop 3 and the corresponding nucleotide sequence based on *E. coli* codon usage.

Red numbers on the above show the order of this AA in the 3rd loop of human RhD protein sequence. Bottom box shows the reverse translation of this AA sequence back to the nucleotide level based on the *E. coli* codon usage table.

1	2	3	4	5	6	7	8	9	10
A	L	L	R	S	P	I	E	R	K
GCG	CTG	CTG	CGT	AGC	CCG	ATC	GAG	CGT	AAG

Figure 3.9. Amino acid sequence of Human RhD external loop 4 and the corresponding nucleotide sequence based on *E. coli* codon usage.

Red numbers on the above show the order of this AA in the 4th loop of human RhD protein sequence. Bottom box shows the reverse translation of this AA sequence back to the nucleotide level based on the *E. coli* codon usage table.

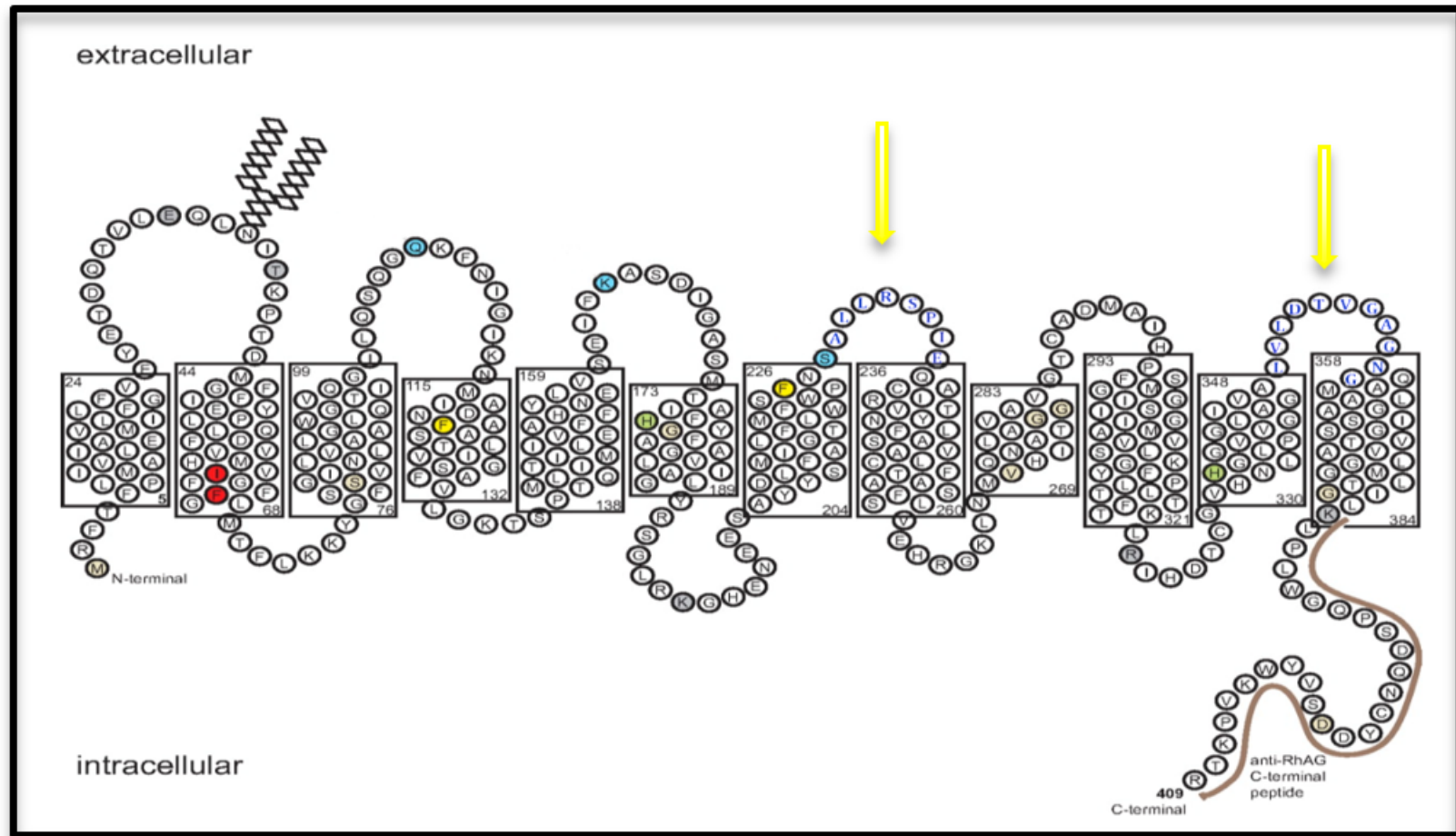


Figure 3.10. Predicted membrane topology and amino acid sequence of *NeRh50*^{4,6} construct.

This figure depicts the hypothetical membrane orientation of the recombinant *NeRh50*^{4,6} in which predicted externalized amino acids are numbered. The recombinant protein was derived initially from a cDNA of a wild type *NeRh50* combined with a human RhD external loops 4 and 6 polypeptide sequences. The predicted extracellular loops of the human RhD amino acids loop 6 [LVLDTVGAGNGM]/ loop 4 [ALLRSPIERK] are shown as solid circles with a blue single letter amino acid code inside each circle. The yellow arrows point to the external loop 4 and loop 6 respectively.

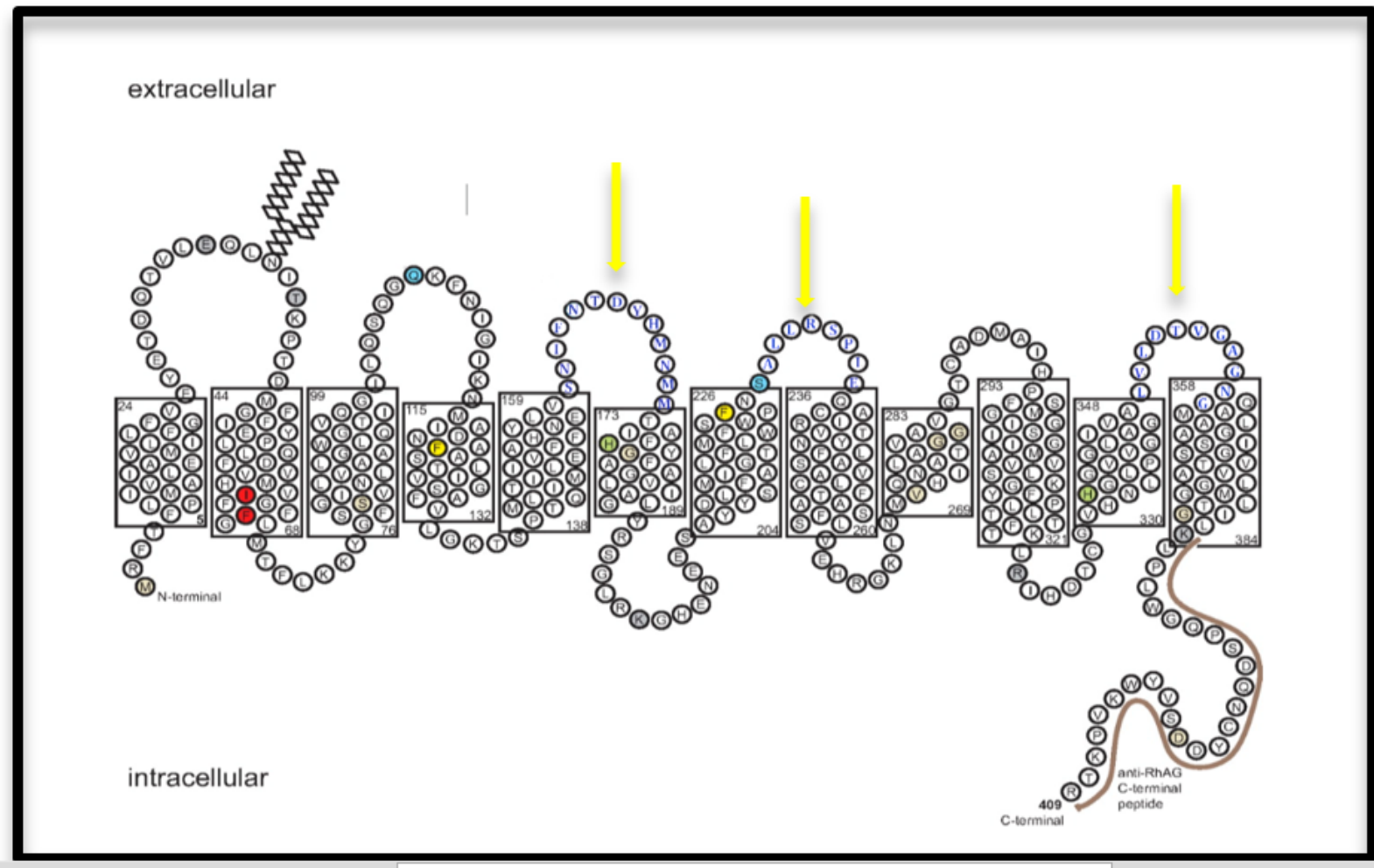


Figure 3.11. Predicted membrane topology and amino acid sequence of *NeRh50*^{3, 4, 6} construct.

This figure depicts the hypothetical membrane orientation of the recombinant *NeRh50*^{3, 4, 6} in which predicted externalized amino acids are numbered. The recombinant protein was derived initially from a cDNA of a wild type *NeRh50* combined with a human RhD external loop 3, loops 4 and 6 polypeptide sequences. The predicted extracellular loops of the human RhD amino acids loop 6 [LVLDTVGAG-NGM]/ loop 4 [ALLRSPIERK]/ loop 3 [NIFNTDYHMMNM] are shown as solid circles with a blue single letter amino acid code inside each circle. The yellow arrows point to the external loops 3, 4 and loop 6 respectively.

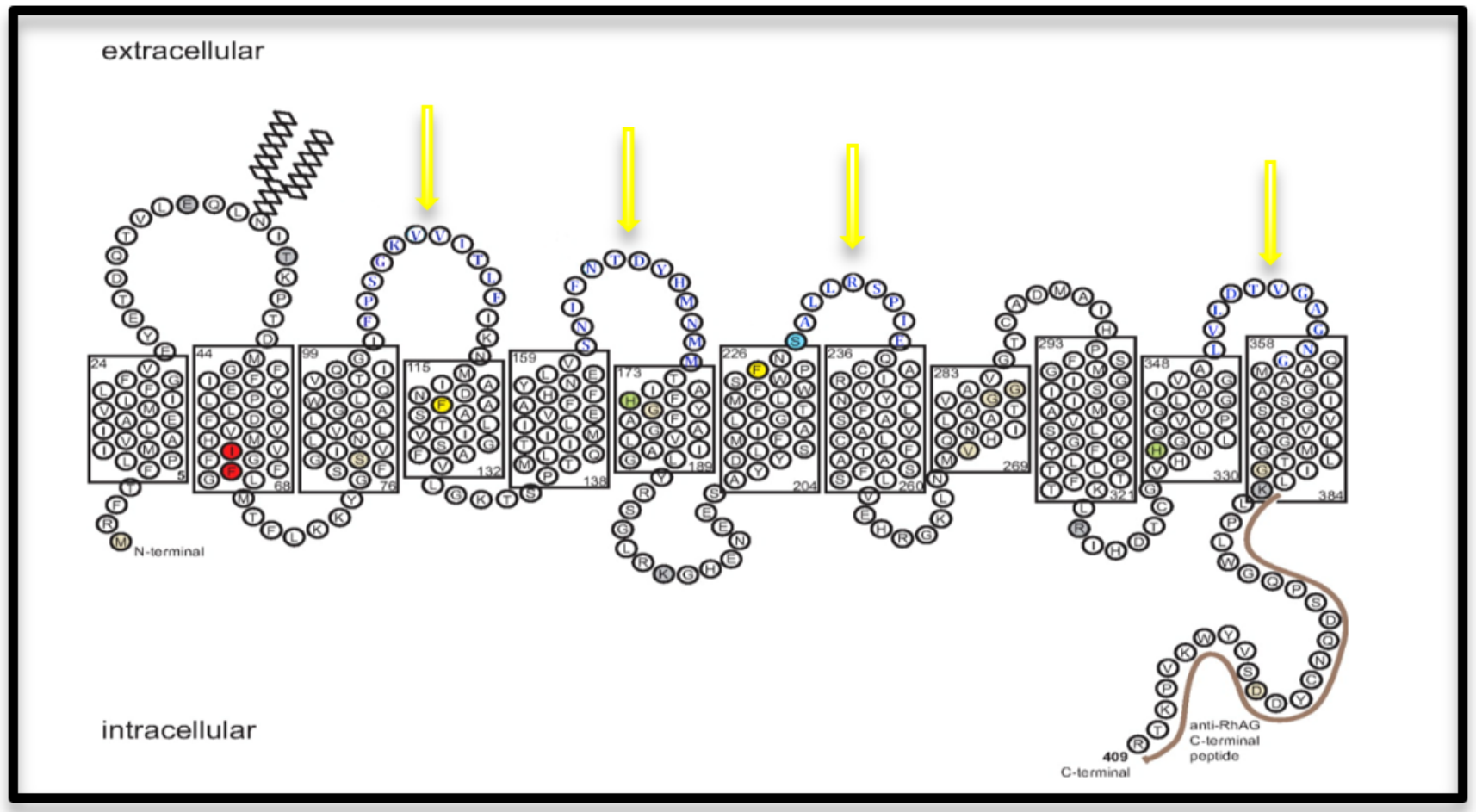


Figure 3.12. Predicted membrane topology and amino acid sequence of *NeRh50*^{2, 3, 4, 6} construct.

This figure depicts the hypothetical membrane orientation of the recombinant *NeRh50*^{2, 3, 4, 6} in which predicted externalized amino acids are numbered. The recombinant protein was derived initially from a cDNA of a wild type *NeRh50* combined with a human RhD external loop 3, loops 4 and 6 polypeptide sequences. The predicted extracellular loops of the human RhD amino acids loop 6 [LVLDTV-GAGNGM]/ loop 4 [ALLRSPIERK]/ loop 3 [NIFNTDYHMMNMM] and loop 2 [FPSGKVVIT] are shown as solid circles with a blue single letter amino acid code inside each circle. The yellow arrows point to the external loops 2, 3, 4 and loop 6 respectively.

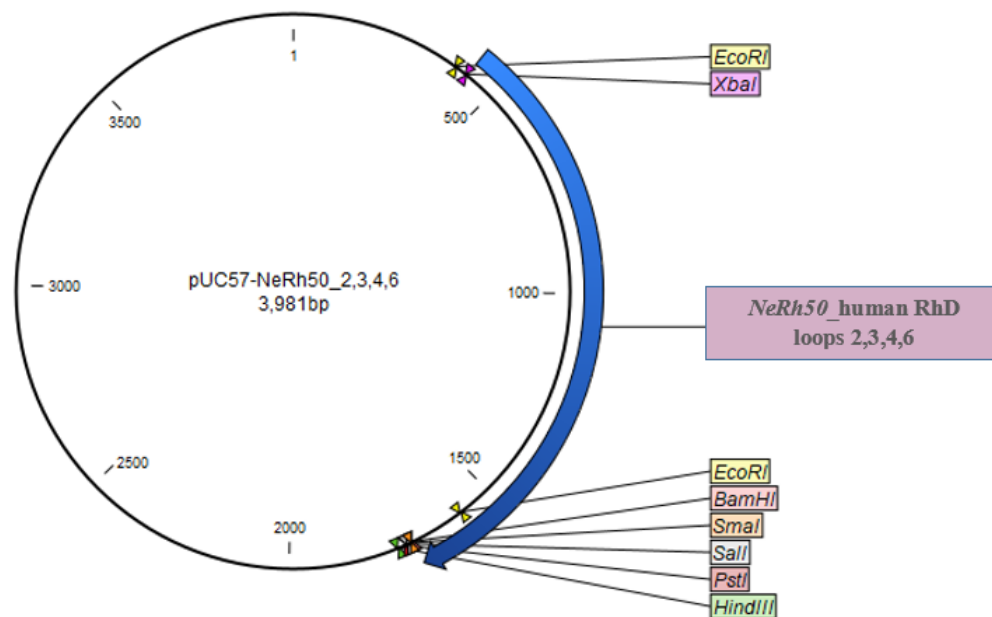


Figure 3.13. pUC57 vector with the recombinant gene (*NeRh50*"2, 3, 4, 6) circular sequence map.

A circular sequence of the full pUC57 vector with the recombinant gene (*NeRh50*"2, 3, 4, 6), it shows all the restriction enzymes sites. By CLC Genomics Workbench 12.0 programme.

3.2.2 Purifying DNA

The pUC57 plasmid with *NeRh50''6* DNA was purified using Wizard^(R) Plus SV Minipreps DNA Purification System from Promega. The DNA concentration was assessed using NanoDrop 2000 UV-Vis Spectrophotometer (Thermos Scientific, UK).

3.2.3 PCR amplifying the *NeRh50''6*

The hybrid recombinant gene (*NeRh50''6*) was amplified using PCR in the presence of AD1 forward and AD2 reverse primers to generate *NeRh50''6** (amplicon) with a new restriction site (*BamHI/ NdeI*) and an amplicon size of 1349 bp. These samples were run on a 1% (w/v) agarose gel to check the size of the recombinant gene *NeRh50''6** (Figure 3.14).

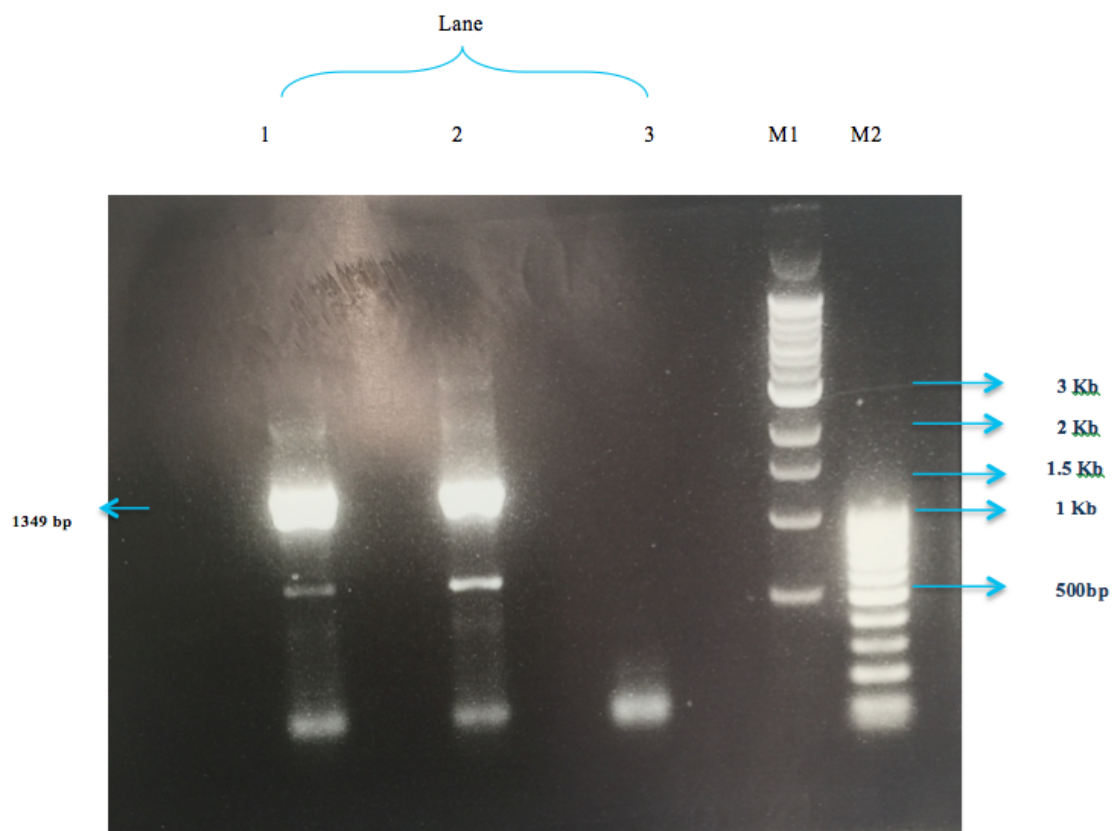


Figure 3.14. Amplicons of pUC57 plasmid - *NeRh50''6**

Lane numbers (1 / 2) are representing 2 amplicons of the pUC57 plasmid - *NeRh50''6**, which were transformed in *DH5 α E. coli*. Using AD1 forward/ AD2 reverse primers formed amplicons; the amplicon size is 1349 bp. Lane number (3) is representing the negative control of the PCR product (nuclease free water). Products were run on a 1% Agarose gel (w/v) alongside DNA markers [M1(1 Kb) and M2(100 bp)]. Shown to the right of the figure are marker sizes in base pairs.

3.3 Expression Vectors

3.3.1 pESV2 expression vector

pESV2 expression vector is the second vector used in this project after pUC57 plasmid because it is useful in downstream techniques to detect the recombinant protein *NeRh50''RhD* human external loops expression on the *E. coli*. pESV2 is an *E. coli* expression vector which was used by Cherif-Zahar *et al.*, 2007, it was used to study the expression of the *NeRh50* protein in *E. coli*.

3.3.1.1 Digestion of pESV2-*NeRh50''6*

Following digestion of the pESV2-*NeRh50''6* plasmid with restriction enzymes *NdeI* and *BamHI* (See 2.2.3.2), the products were run on a 1% (w/v) agarose gel and two bands were seen. The lower band (1.7 kb) represents the digested insertion *NeRh50''6* and the upper band 3 kb represents the pESV2 plasmid (Figure 3.15).

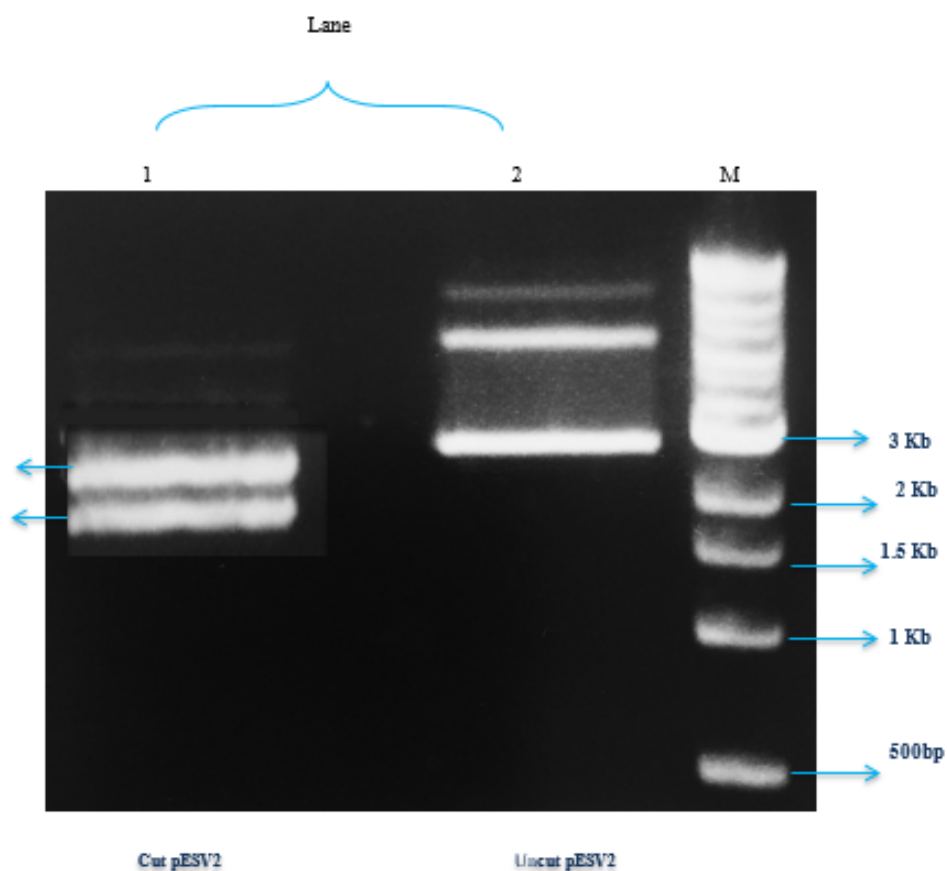


Figure 3.15. *BamHI* / *NdeI* digestion of pESV2- *NeRh50''6*.

Lane (1) the digested pESV2 plasmid, which shows 2 bands where the 2 restriction enzymes (*BamHI* and *NdeI*) cut the plasmid. Bands size are 3 Kb represent the digested pESV2 plasmid and 1.7 Kb represents the insertion. Lane (2) the non-digested plasmid, which is the untreated sample with the (*BamHI* and *NdeI*) restriction enzymes shows many bands due to the multiple constructing processes that the pESV2 vector has been through (see 2.2.2.3.2). Products were run on a 1% (w/v) agarose gel alongside a DNA marker (M). Shown to the right of the figure are marker sizes in base pairs.

3.3.1.2 Cloning of pESV2-*NeRh50''6**[^] Human *RHD* in *DH5α E. coli* Competent Cells

Following the ligation of the mixture of purified pESV2 with the purified digested amplicon (*NeRh50''6**[^]) with two different ratio (see 2.2.3.4), the product was transformed in *DH5α E. coli* competent cells. After that, master plate was prepared. PCR screening for single colonies with different ligation ratio were run on a 1% (w/v) agarose gel. The 2:1 ratio (Figure 3.16) (A) showed a single band with high intensity a round 1.3 kb in all colonies, which represent the right size of *NeRh50''6**[^] in pESV2 expression vector after amplified with AD1 / AD2 primers. The 5:1 ratio (B) showed lower band intensity than 2:1 and in some colonies it showed multiply bands.

The PCR screening was done to check if each colony contained the positive plasmid, which has the correct insertion while the master plate was done to keep a stock from each positive colony for later use.

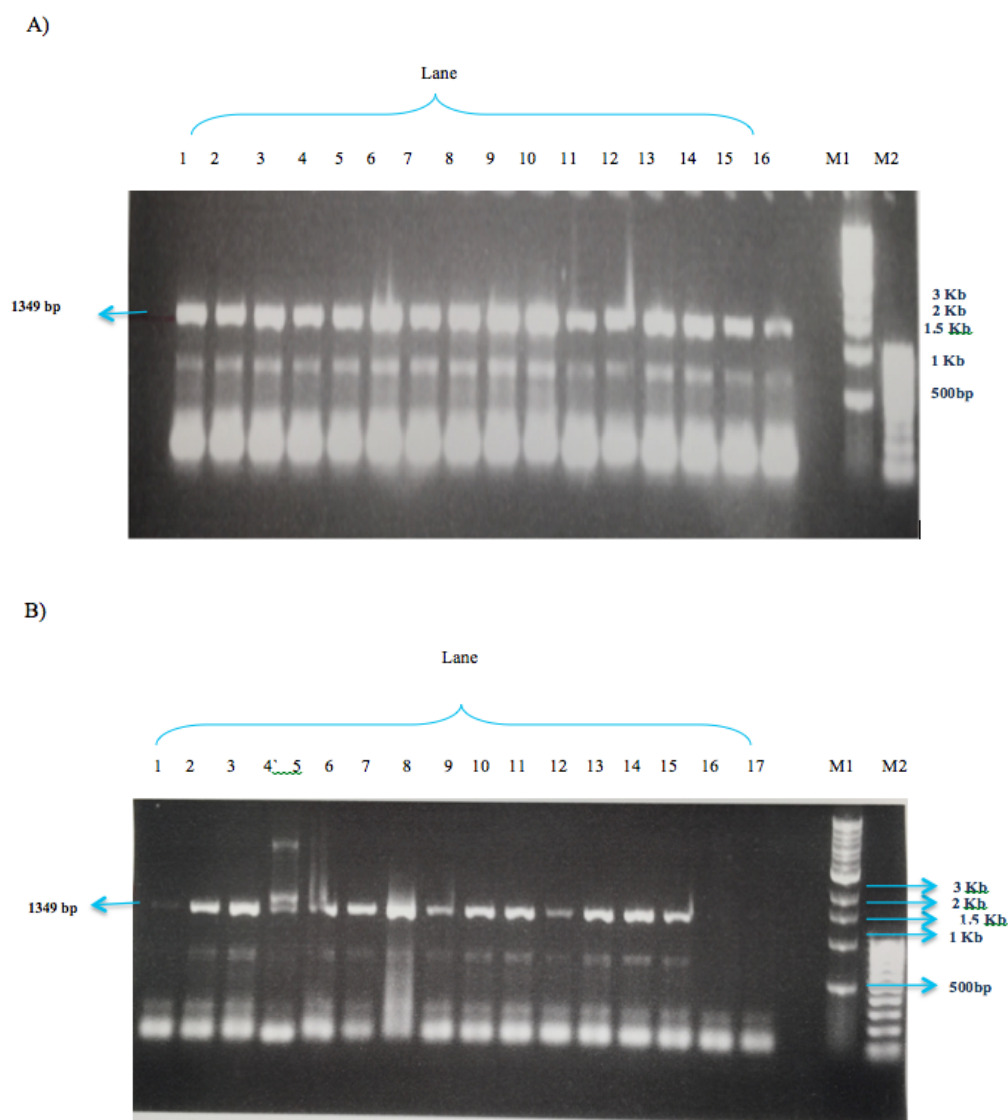


Figure 3.16. PCR screening of positive pESV2-*NeRh50''6** plasmid colonies, which amplified with AD1 / AD2 primers.

Amplicon samples of the single colonies: from the transformation plate of the [pESV2 with cDNA (*NeRh50''6*)] with *DH5 α E. coli* competent cells with different ligation ratio. A) Lane (1-16) shows amplicon samples of [pESV2 with cDNA (*NeRh50''6**)] transformed with *DH5 α E. coli* competent cells with (2:1) ligation ratio insert: vector. B) Lane (1-17) shows amplicon samples of [pESV2 with cDNA (*NeRh50''6**)] transformed with *DH5 α E. coli* competent cells with (5:1) ligation ratio insert:vector. Products were run on a 1% (w/v) Agarose gel alongside DNA markers (M1 and M2). Shown to the right of figure are marker sizes in base pairs.

3.3.1.3 Sequencing of pESV2 expression vector

pESV2 vector was constructed from pACYC184 and pSU18 vectors and it was combined with sequence from other vectors (Figure 3.18) as described in Javelle *et al.*, 2003.

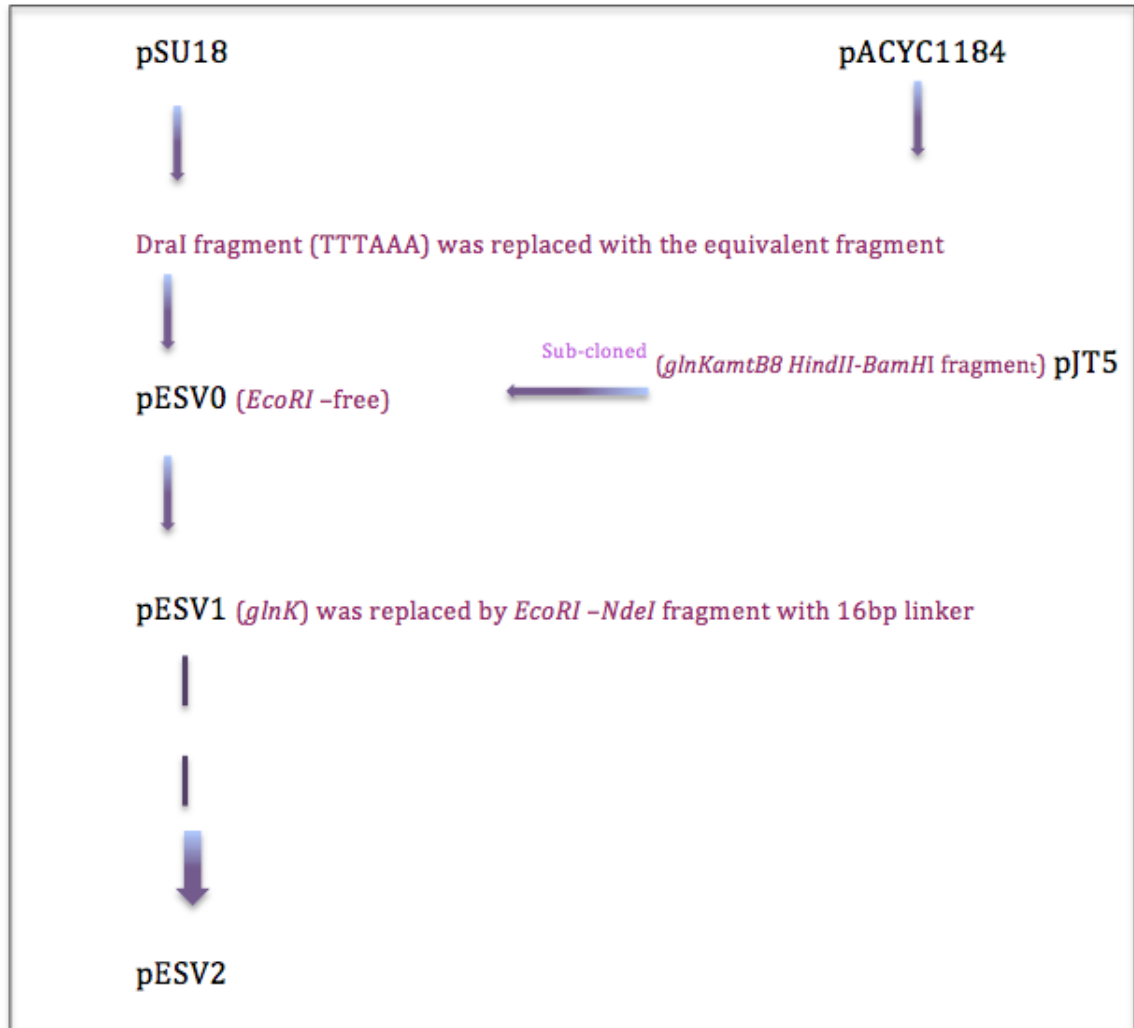


Figure 3.17. pESV2 vector constructed process.
Javelle *et al.*, 2003

Snap Gene program was used to produce a pESV0 vector map, which is part of this process of constructing the pESV2. This vector contains the AmtB Shine Dalgarno sequence (5-AGGAGGU-3), which is a (6-8) base sequence located upstream of the start codon of the bacterial mRNA, that has an importance in ribosome alignment for efficient translation. Moreover, other critical sequences were placed in the promoter sequence of the vector to make sure that the vector will work in *E. coli*.

The problem was the unavailability of either the vector map or the sequence of some vectors (provided by Prof Mike Merrick), which were used to construct the pESV2 vector. Therefore, we sequenced one of the original samples from Prof. Merrick's lab (Figure 3.18) and one of the clones, which has the vector from Prof. Merrick's lab combined with *NeRh50''6** (Figure 3.19) by next generation sequencing (See 2.2.2.3.7.3). So, a total of 2 samples, were sequenced using Ion PGMTM in one run. The data was then processed, and a summary was produced by the Torrent SuiteTM Software Version 4.4 (Figure 3.18 and 3.19).

The total reads produced from the original sample was 460,133, while the total reads from the vector combined with *NeRh50''6** was 397,704. In the first report (Figure 3.18), the Ion Sphere Particle (ISP) loading density was 90%, while 10% of the 314R chip wells were empty. Moreover, in the second report (Figure 3.19), the Ion Sphere Particle (ISP) loading density was 85%, while 15% of the 314R chip wells were empty. A brief sequencing

report of the original sample sequencing run is displayed in Figure 3.18. For downstream analysis, there were around 460,133 usable reads in total, which is 41% of the total reads. These reads were thoroughly processed by well classification and read filtering to ensure their quality. Clonal ISPs (in which single template fragments were amplified) made up 52% of the reads, while 48% were polyclonal ISPs. After filtering out 1% of test fragments and 22% of low-quality reads, the final percentage of the final library was 77% with a mean read length of 180 bp (Figure 3.18). Furthermore, Figure 3.19 shows a brief sequencing report for the sample containing the vector combined with *NeRh50"6**. For downstream analysis, there were around 397,704 usable reads in total, which is 37% of the total reads. These reads were thoroughly processed by well classification and read filtering to ensure their quality. Clonal ISPs (in which single template fragments were amplified) was made up 51% of reads, while 49% were polyclonal ISPs. After filtering out 1% of test fragments and 27% of low-quality reads, the final percentage of the final library was 72% with a mean read length of 177 bp (Figure 3.19).

The outcome of the NGS work confirmed that; the original vector from Prof. Merrick's lab was, pUC57 vector after the sequence were Blasted to vector databases (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome) because there were no ref-

erence sequence to compare it with (see appendix A). The sequence from our NGS data (100%) similarity alignment of the nucleotides to the pUC57 vector sequence, provided by GenScript. Moreover, appendix B shows the alignment of the *NeRh50"6** sequence in pESV2 vector provided from Prof. Merrick's lab to the *NeRh50"6* sequence in pUC57 vector, provided by GenScript.

Run Summary

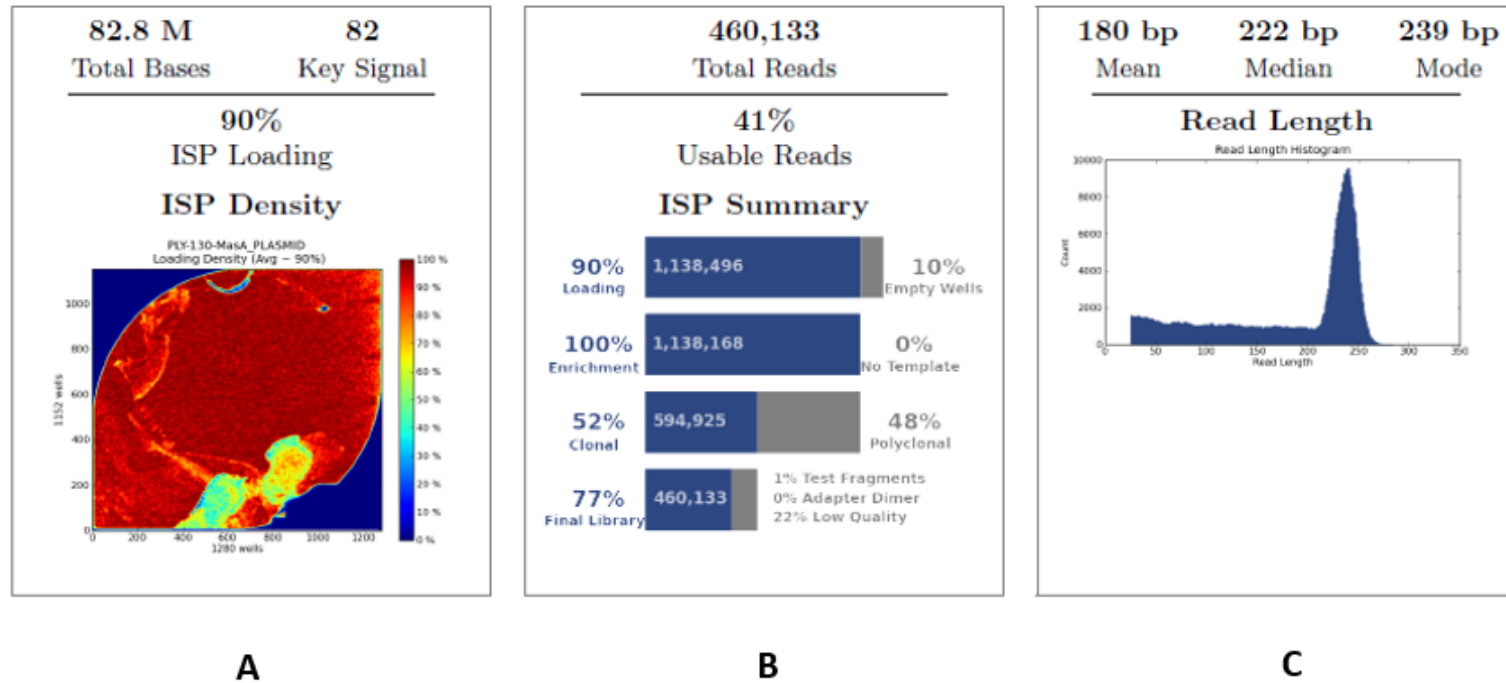


Figure 3.18. The report of single sequencing run of the original (pESV2) vector sample library

A) The percentage of ISP loading density of the ISP shown by the 314R chip wells was 90%. Different colours represent the loading percentage of ISP through the physical 314R chip plate surface (red is highest; blue is lowest).

B) The total number of usable reads is 460,133, after trimming and filtration from empty wells, non-templated and polyclonal reads. The percentage 41% is obtained by dividing these reads by the number of reads containing the library ISPs (1,138,496). The live/enrichment percentage is 100%, which indicates that ISPs contain a strong sequence signal from test fragment and library (templated).

C) Histogram shows a mean reading length of 180 bp. The read count is displayed in the y-axis, while the read length, in bp, is shown on the x-axis.

Run Summary

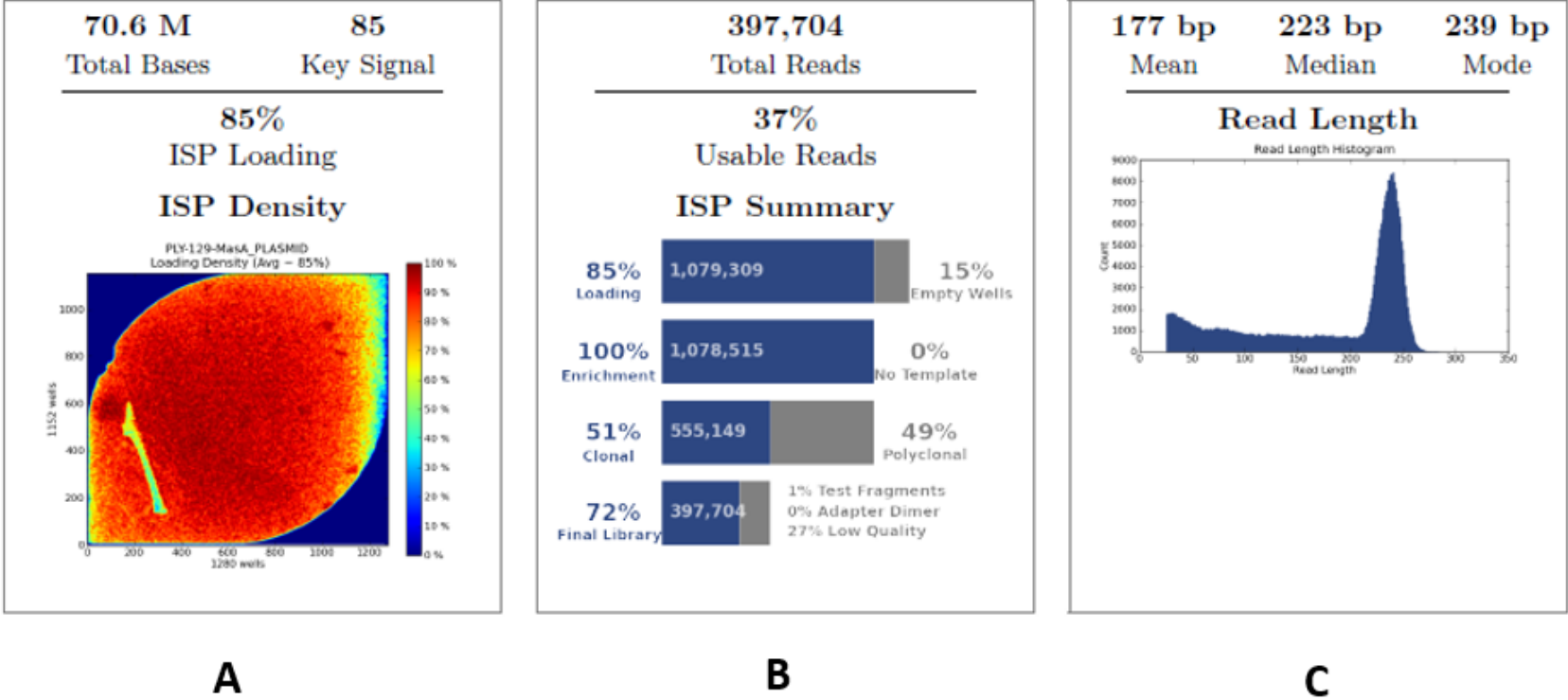


Figure 3.19. The report of single sequencing run of the pESV2- *NeRh50"6* sample library

- A) The percentage of ISP loading density of the ISP shown by the 314R chip wells was 85%. Different colours represent the loading percentage of ISP through the physical 314R chip plate surface (red is highest; blue is lowest).
- B) The total number of usable reads is 397,704, after trimming and filtration from empty wells, non-templated and polyclonal reads. The percentage 37% is obtained by dividing these reads by the number of reads containing the library ISPs (1,079,309). The live/enrichment percentage is 100%, which indicates that ISPs contain a strong sequence signal from test fragment and library (templated).
- C) Histogram shows a mean reading length of 177 bp. The read count is displayed in the y-axis, while the read length, in bp, is shown on the x-axis.

Therefore, due to the sequencing results of the vector provided from Merrick's lab, which came as a pUC57 vector instead of pESV2 vector, it was no longer used and it was replaced with pET-a9 expression vector, which is widely used for expression work.

3.3.2 pET-a9 expression vector

pET-a9 expression vector (Figure 3.20) is a vector, which carries an N-terminal T7●Tag[®] sequence and *Bam*HI cloning site. It is one of the best system for cloning and expressing recombinant proteins in *E. coli*.

Figure has been removed due to Copyright restrictions.

Figure 3.20. pET-9a expression vector map showing all restriction enzyme sites
adapted from Novagen

3.3.2.1 Cloning of pET- *NeRh50*/Human RhD external loops in BL21 (DE3) Chemically Competent Cells

BL21 (DE3) competent cells are chemically competent *E. coli* cells used for high-level protein expression with T7 RNA polymerase-based expression systems, which is a derivative of T7 expression *E. coli* strain.

The mixture of purified digested pET with the purified digested *NeRh50* human RhD external loops from ligation step (see 2.2.2.3.4) was transformed in BL21 (DE3) *E. coli* chemically competent cells; it was incubated for 16 hours at 37 °C. After that, master plate was prepared. Thereafter, an enzymatic digestion with *BamHI* and *XbaI* step was run on each single colony from the master plate to check the positive colony, which has the correct insertion (*NeRh50* human RhD external loops); *NeRh50*6 with 1766 bp, *NeRh50*4, 6 with 1275bp, *NeRh50*3, 4, 6 with 1275 bp, and *NeRh50*2, 3, 4, 6 with 1275 bp, as shown in Figure 3.21. Therefore, according to the band size (Figure 3.21) prove that the ligation step was successful and the insertion of *NeRh50*RHD with different constructs were successfully cloned in pET expression vector.

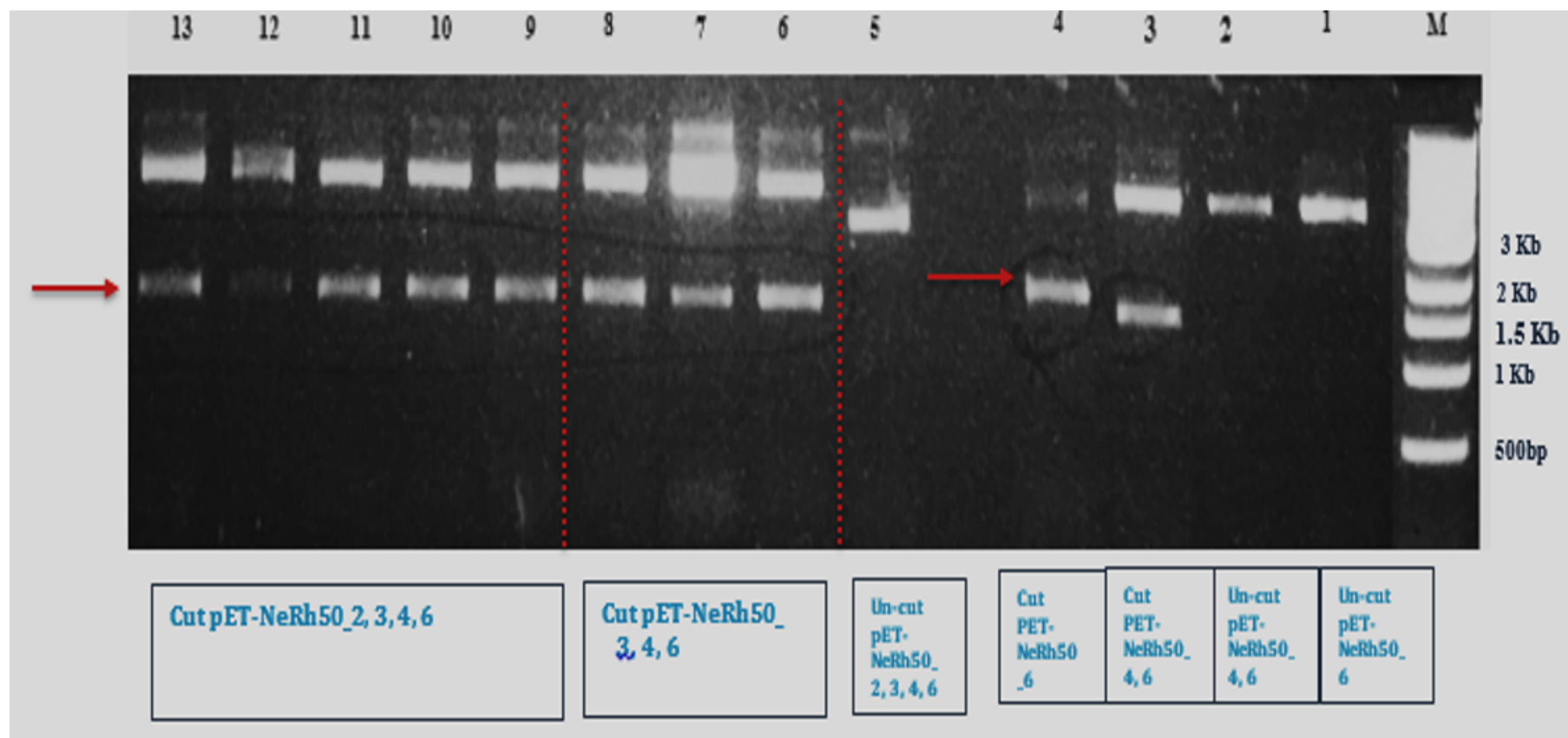


Figure 3.21. Digested positive colonies of pET vector with the *NeRh50*"6, *NeRh50*"4, 6, *NeRh50*"3, 4, 6, *NeRh50*"2, 3, 4, 6 constructs.

Lane (1, 2) non-digested pET- *NeRh50*"6 gene and pET- *NeRh50*"4, 6 gene respectively, shows band around 5Kb. Lane(3) Digested pET- *NeRh50*"4, 6", which shows 2 bands where the 2 restriction enzymes (*Bam*HI and *Xba*I) cut the plasmid. Bands size around 4 Kb and 1275 bp. Lane(4) Digested pET- *NeRh50*"6", which shows 2 bands where the 2 restriction enzymes (*Bam*HI and *Xba*I) cut the plasmid. Bands size around 4 Kb and 1766 bp. Lane (5) non-digested pET- *NeRh50*"2, 3, 4, 6 gene, shows band around 5Kb. Lane(6-8) Digested pET-*NeRh50*"3, 4, 6", which shows 2 bands where the 2 restriction enzymes (*Bam*HI and *Xba*I) cut the plasmid. Lane(9-13) Digested pET-*NeRh50*"2, 3, 4, 6", which shows 2 bands where the 2 restriction enzymes (*Bam*HI and *Xba*I). Bands size around 4 Kb and 1275 bp. Products were run on a 1% (w/v) agarose gel alongside a DNA marker (M). Shown to the right of figure are marker sizes in base pairs. Arrows point to the band of interest.

3.3.2.2 Sanger Sequencing of *NeRh50*"human *RhD* external loops

to confirm the presence of the human RhD external loops Sanger sequencing was used to sequence cDNA clones of each construct: *NeRh50*"6, *NeRh50*"4, 6, *NeRh50*"3, 4, 6, and *NeRh50*"2, 3, 4, 6, to confirm restriction digestion results (Figure 3.21) using specific primer sets (Table 2.4). Primers were designed to amplify the sequences around each human RhD external loops, moreover, there were two primer sets to amplify the region before and after the hybrid sequence *NeRh50*" human RhD external loops in the pET expression vector sequence (Figure 3.22). The only primer sets that worked were the sets around the hybrid gene in the pET expression vector sequence (Figures 3.23, 3.24, 3.25, and 3.26), while other primer sets gave unusable sequence or some of them did not amplify at all (see Appendices C-I for full alignment of all the constructs against each primer). However, Sanger sequencing data were aligned against pET- *NeRh50*"2, 3, 4, 6 sequence. F1 forward primer (Figure 3.23) and R1 reverse primer (Figure 3.24) which aligned with the sequence before the hybrid gene by using CLC Genomics Workbench 12.0 programme. Moreover, F7 forward primer (Figure 3.25) and R7 reverse primer (Figure 3.26) were used to amplify the region on the pET expression vector after the insertion location. The sequence data confirmed that we have the pET expression vector and the cloning process was successful, though the coverage of these primers (F1, F7, R1, and R7) was not enough to read

entire hybrid recombinant sequence.

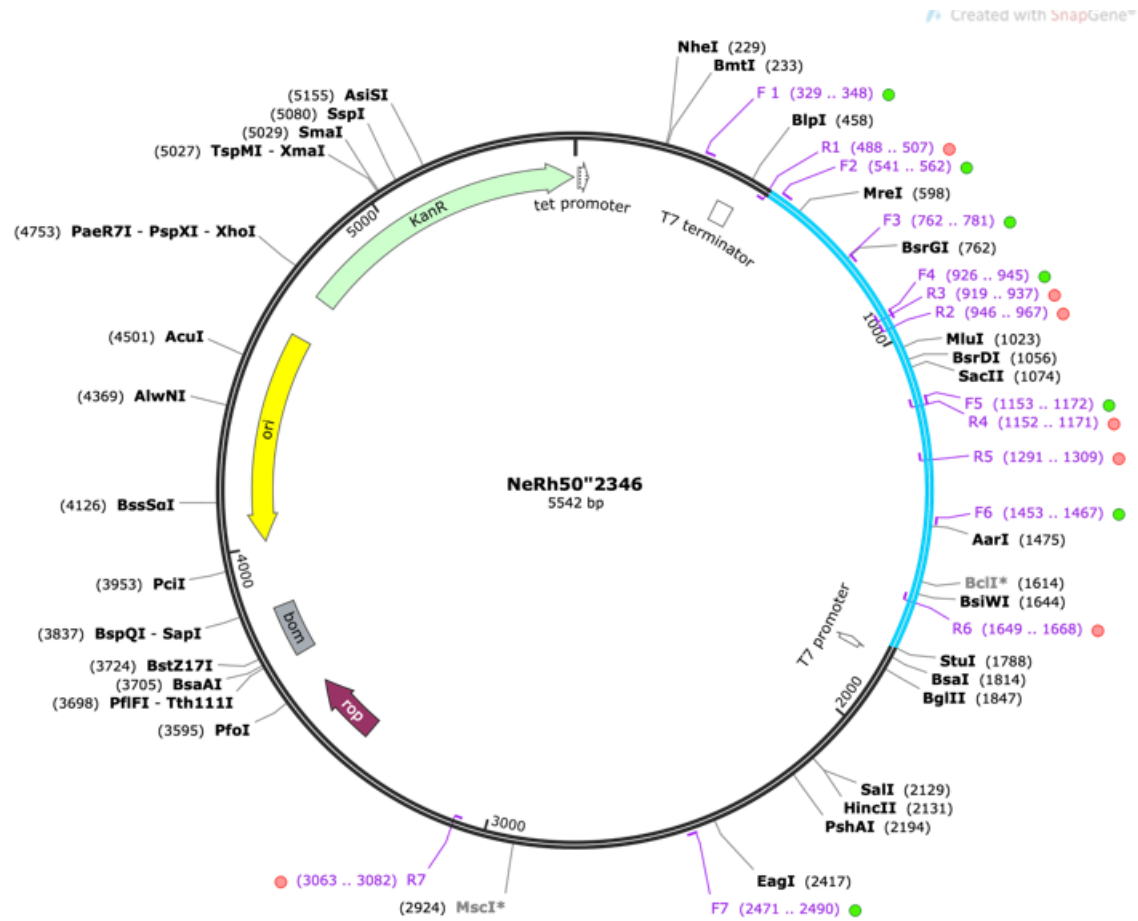


Figure 3.22. pET-NeRh50"2,3,4,6 vector map

pET-NeRh50"2,3,4,6 vector map showing all restriction enzyme sites and all primer sets(F1-F7) with green dots, (R1-R7) with red dots. Blue part shows recombinant *NeRh50"2,3,4,6* DNA. by *SnapGene*.

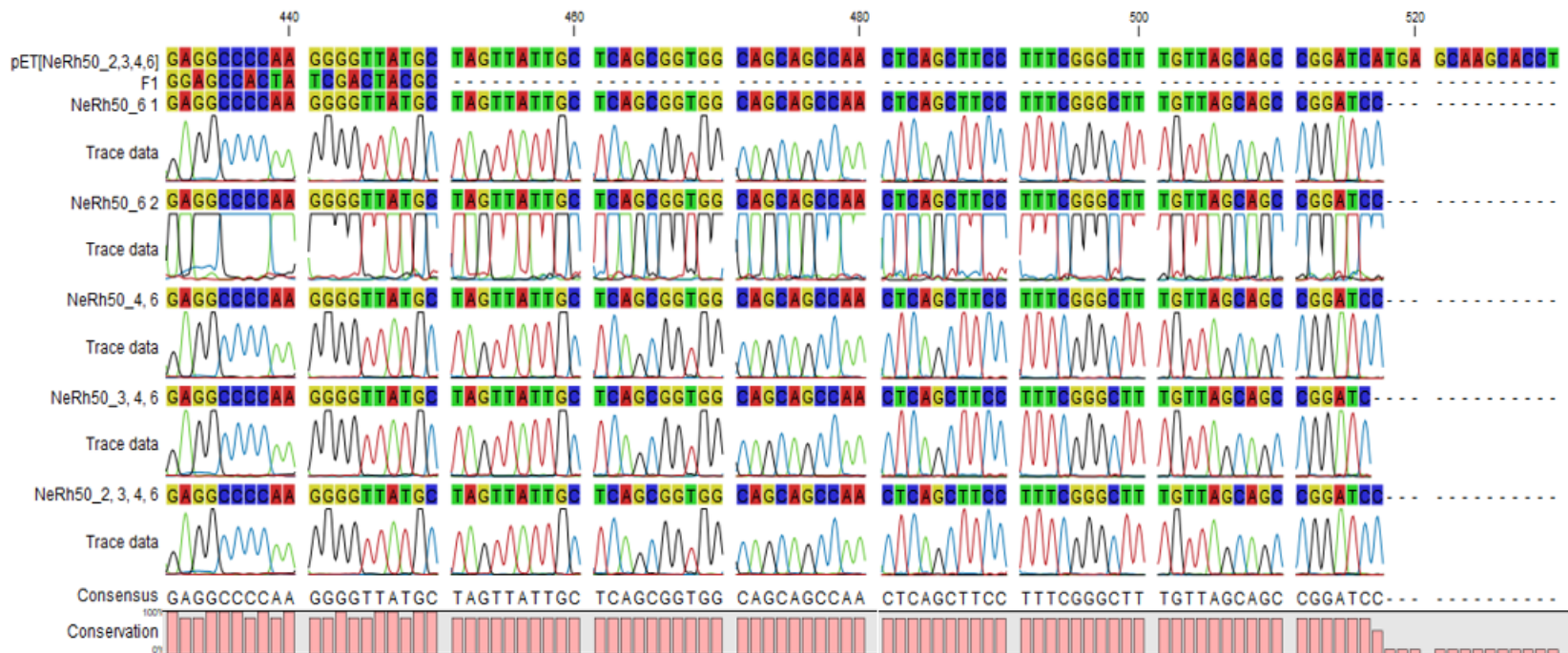


Figure 3.23. Sanger sequencing data with electropherogram of all constructs cDNAs against (F1) forward primer.

Sanger sequencing data for all constructs: *NeRh50*"6, *NeRh50*"4, 6, *NeRh50*"3, 4, 6, and *NeRh50*"2, 3, 4, 6 were aligned by CLC Genomics Workbench 12.0 programme against: F1 forward primer which cover the sequence in the pET expression vector before insertion. Moreover, that were aligned against the reference sequence, which is pET-*NeRh50*"2, 3, 4, 6.

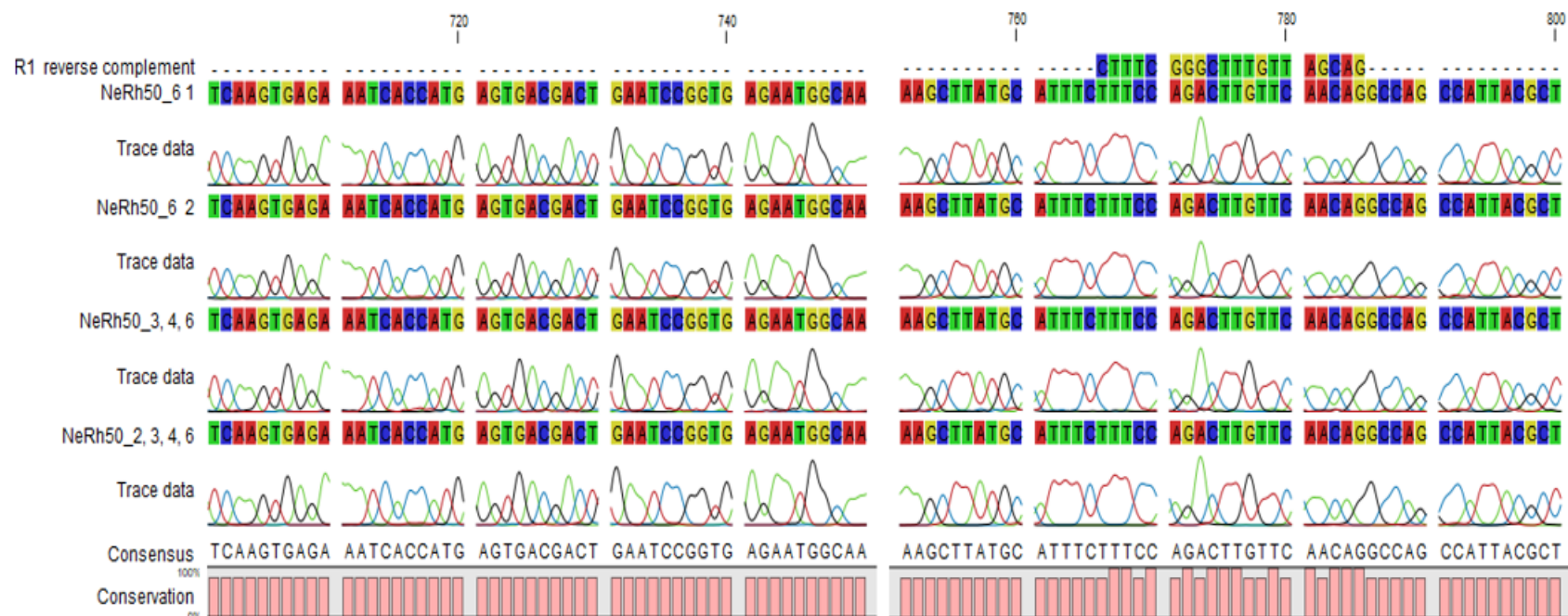


Figure 3.24. Sanger sequencing data with electropherogram of all constructs cDNAs against (R1) reverse primer.

Sanger sequencing data for all constructs: *NeRh50*"6, *NeRh50*"4, 6, *NeRh50*"3, 4, 6, and *NeRh50*"2, 3, 4, 6 were aligned by CLC Genomics Workbench 12.0 programme against: R1 reverse primer, which cover the sequence in the pET expression vector before insertion.

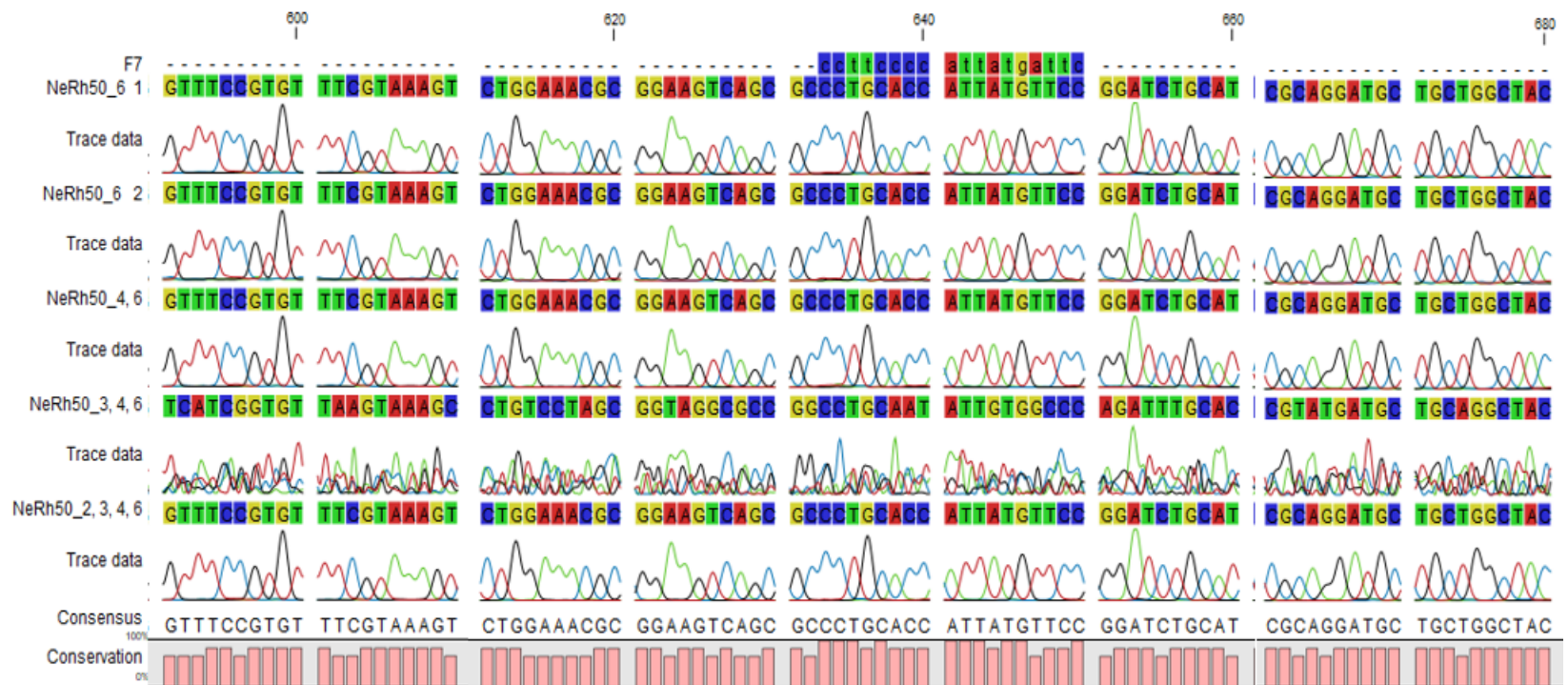


Figure 3.25. Sanger sequencing data with electropherogram of all constructs cDNAs against (F7) forward primer.

Sanger sequencing data for all constructs: *NeRh50*"6, *NeRh50*"4, 6, *NeRh50*"3, 4, 6, and *NeRh50*"2, 3, 4, 6 were aligned by CLC Genomics Workbench 12.0 programme against: F7 forward primer which cover the sequence in the pET expression vector after insertion.

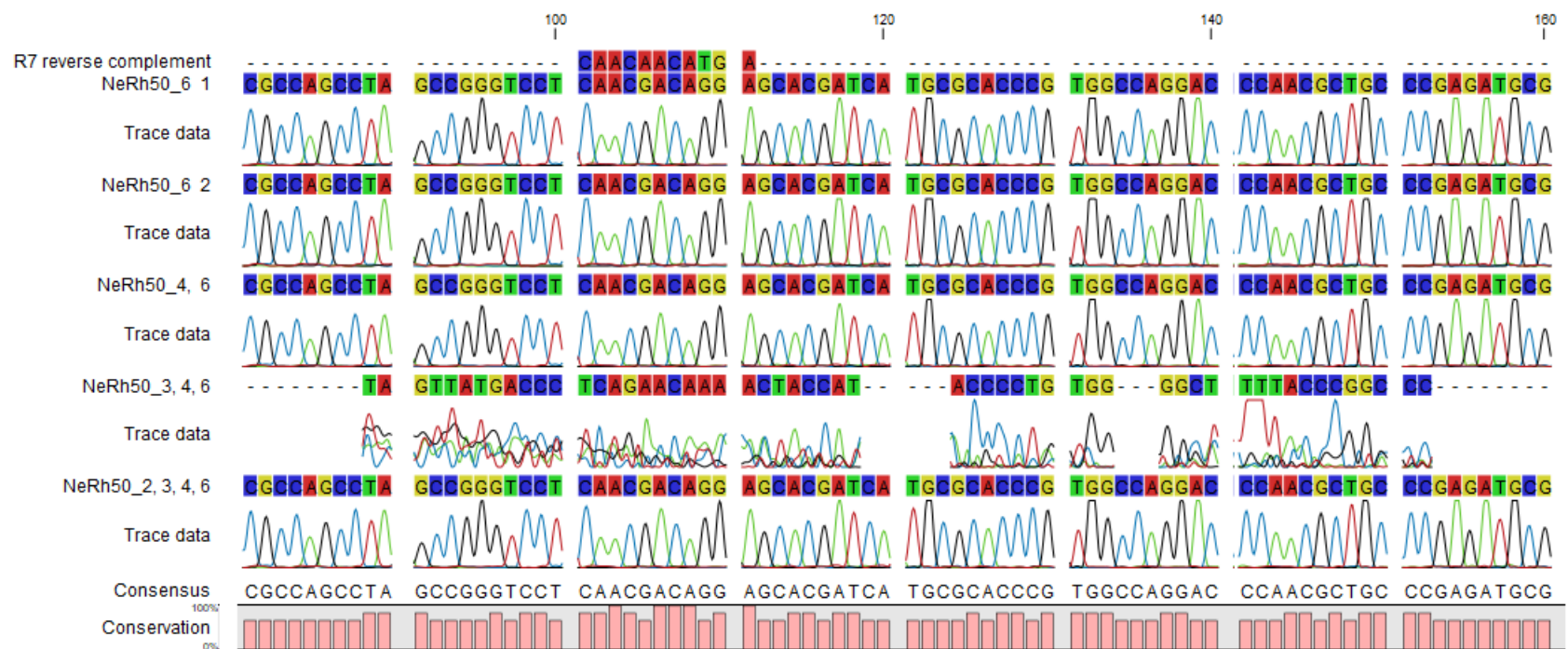


Figure 3.26. Sanger sequencing data with electropherogram of all constructs cDNAs against (R7) forward primer.

Sanger sequencing data for all constructs: *NeRh50*"6, *NeRh50*"4, 6, *NeRh50*"3, 4, 6, and *NeRh50*"2, 3, 4, 6 were aligned by CLC Genomics Workbench 12.0 programme against: R7 forward primer which cover the sequence in the pET expression vector after insertion.

3.4 Three-Dimensional Protein Model

SWISS-MODEL is a fully automated protein structure homology-modelling server. The purpose of this server is to create protein models (<https://swissmodel.expasy.org/>).

This program was used to create the 3-dimensional protein model for construct; *NeRh50*"2, 3, 4, 6 (Figure 3.30). Moreover, the software tool was used to compare the three-dimensional protein module for the construct to the human RhD wild type protein (Figure 3.28) and *NeRh50* wild type protein (Figure 3.29). The construct was compared to the external loops of the human RhD wild type protein and to the backbone of *NeRh50* wild type protein in order detecting changes in these aspects.

In the *NeRh50*"2, 3, 4, 6 (Figure 3.27) there are light blue circles which are present in the *NeRh50* wild type protein (Figure 3.29) that represent the backbone of the construct. Moreover, the dark blue circles represent human RhD external loop 2, orange circles represent human RhD external loop 3, green represent human RhD external loop 4, and cyan represent human RhD external loop 6 that have the same colour as the external loops in the human RhD wild type (Figure 3.28).

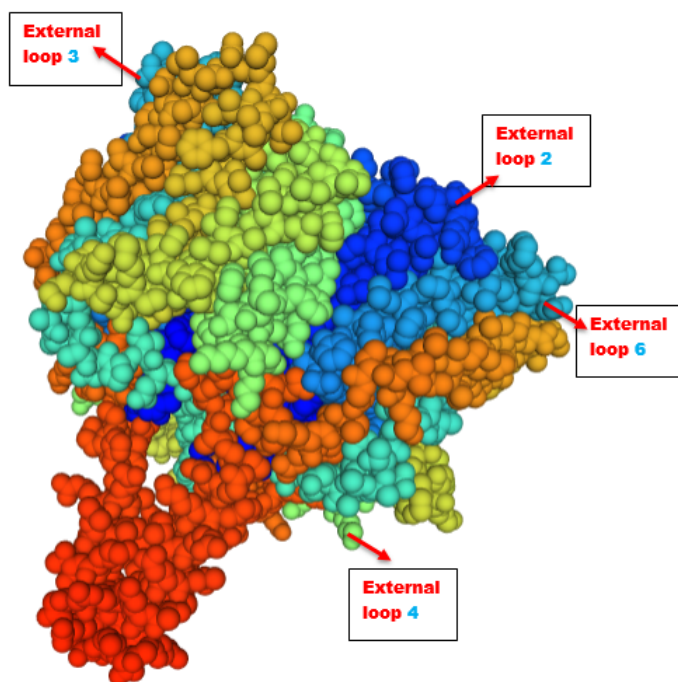


Figure 3.27. Predicted 3D protein model of *NeRh50*" 2, 3, 4, 6.

Predicted 3D protein model of *NeRh50*"2, 3, 4, 6. The colour (dark blue light blue green cyan orange red) represents the amino acid sequence. The arrows show the human RhD external loops (2, 3, 4, 6). Dark blue represents human RhD external loop 2, cyan represents human RhD external loop 6. Orange represents human RhD external loop 3, and green represents human RhD external loop 4. Arrows show external loops of the human RhD. By *SWISS-MODEL*

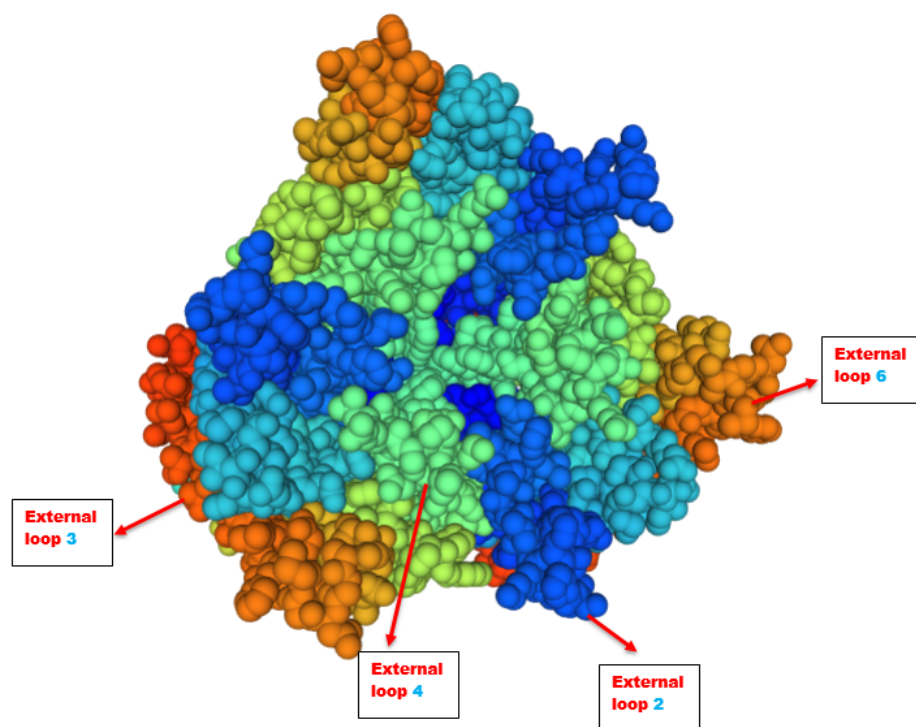


Figure 3.28. Predicted 3D protein model of the human RhD wild type

Predicted 3D protein model of human RhD wild type. The colour; (dark blue light blue green cyan orange red) represents the amino acid sequence. The arrows show the human RhD external loops (2, 3, 4, 6). Dark blue represents human RhD external loop 2, cyan represents human RhD external loop 6. Orange represents human RhD external loop 3, and green represents human RhD external loop 4. Arrows show external loops of the human RhD. By *SWISS-MODEL*

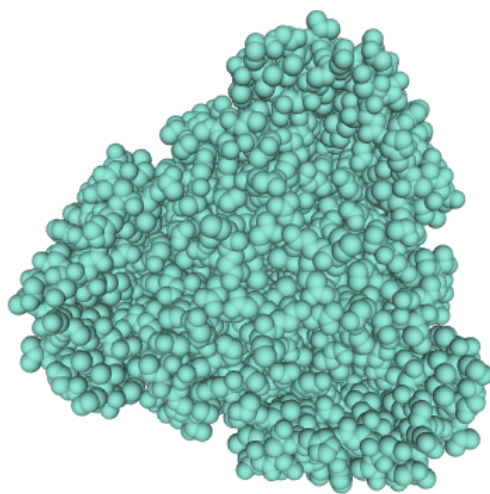


Figure 3.29. Predicted 3D protein model of *NeRh50* wild type

By *SWISS-MODEL*

3.5 Confirming the presence of the Rh Recombinant proteins

Antibody	Western Blotting	Flow cytometry	TEM
2.2B9	*	*	*
2A4	*	*	*
5G9	*	*	
BIRMA-D6	*	*	
MS-26		*	
BIRMA-D10		*	
RUM-1		*	
HAM-A		*	
MAD-2		*	
BRAD-3		*	
1917	*		
1918	*	*	*

Table 3.2. Anti-D and Anti-*NeRh50* antibodies list against different techniques.

List shows Anti-D and Anti-*NeRh50* antibodies, which were used in western blot, flow-cytometry and TEM techniques. Asterisk mark used to show each antibody were used.

3.5.1 Western Blot Results

Seven antibodies (Table 2.1, 3.2) were tested against ghost membrane samples with (Rh⁺, Rh⁻) phenotypes, and the four different constructs expressed in *E. coli*; *NeRh50*"6, *NeRh50*"4, 6, *NeRh50*"3, 4, 6, and *NeRh50*"2, 3, 4, 6, human RhD30 wild type and *NeRh50* wild type to confirm that these antibodies bind to the human RhD external loops and *N.europaea* external loop 1 and C-terminus by western blot.

3.5.1.1 Comparison of Bacterial Protein Extraction Methods

Two different buffers were used; the lysis buffer and BugBuster master mix protein extraction reagent (see 2.2.2.4.7.1) to assess which resulted in a higher protein concentration for the samples. Following use of the Qubit assay for protein quantitation of these samples, it was shown that use of BugBuster master mix protein extraction reagent to have a higher protein concentration in the cell lysates (Table 3.3), also easier and faster than the lysis buffer.

Following this comparison, lysis buffer was no longer used and bacterial proteins for samples from all constructs were extracted by BugBuster master mix protein extraction reagent.

Sample name	Protein concentration following use of lysis buffer $\mu\text{g}/\text{ml}$	Protein concentration following use of Bug Buster $\mu\text{g}/\text{ml}$
<i>NeRh50_6 1</i> (A)	4.53	7.5
<i>NeRh50_6 1</i> (B)	4.35	9.52
<i>NeRh50_6 3</i> (A)	7.6	12.5
<i>NeRh50_6 3</i> (B)	3.31	5.42
<i>NeRh50_6 5</i> (A)	2.33	5.64
<i>NeRh50_6 5</i> (B)	1.69	3.23

Table 3.3. Qubit protein quantitation results of three different colonies of pESV2-*NeRh50"6*

Qubit results of pESV2-*NeRh50"6* colonies after lysing them with lysis buffer or Bug Buster protein extraction reagent. Protein concentrations were higher with the Bug-Buster Master Mix than with the lysis buffer. Moreover, A tube (the pellet samples) has shown much higher results than B tubes (supernatant samples) (see 2.2.2.4.7.1)

3.5.1.2 Detection of human RhD external loop 4 by 2.2B9 antibody by western blotting

2.2B9 a monoclonal antibody (see 2.1.1) was used to detect the expression of human RhD external loop 4 in all constructs and control samples. It was raised against a synthetic peptide corresponding to the fourth external domain of the RhD protein (Figure 3.30) (Skinner, Mushens & Avent, unpublished).

NeRh50 wild type, which was used as a negative control for 2.2B9, shows an immunostaining band close to 37 kDa.

A ghost membrane sample with an Rh⁻ phenotype, was used as a second negative control to 2.2B9, and shows only one faint band around 34 kDa.

On the other hand, a ghost membrane sample with an Rh⁺ phenotype, was used as a positive control to this antibody, and shows clear double band around 32-34 kDa. Moreover, human RhD30 wild type sample shows a clear double band around 32-34 kDa similar to the ghost membrane sample with Rh⁺ phenotype.

pET*NeRh50*"6 sample, shows only one faint band around 37 kDa compared to other constructs; pET *NeRh50*"4, 6, pET *NeRh50*"3, 4, 6, and pET *NeRh50*"2, 3, 4, 6 all react similarly and show a clear double band around 37-36 kDa.

The lower band, which is the band of interest has a band size approximately 36 kDa (Figure 3.30).

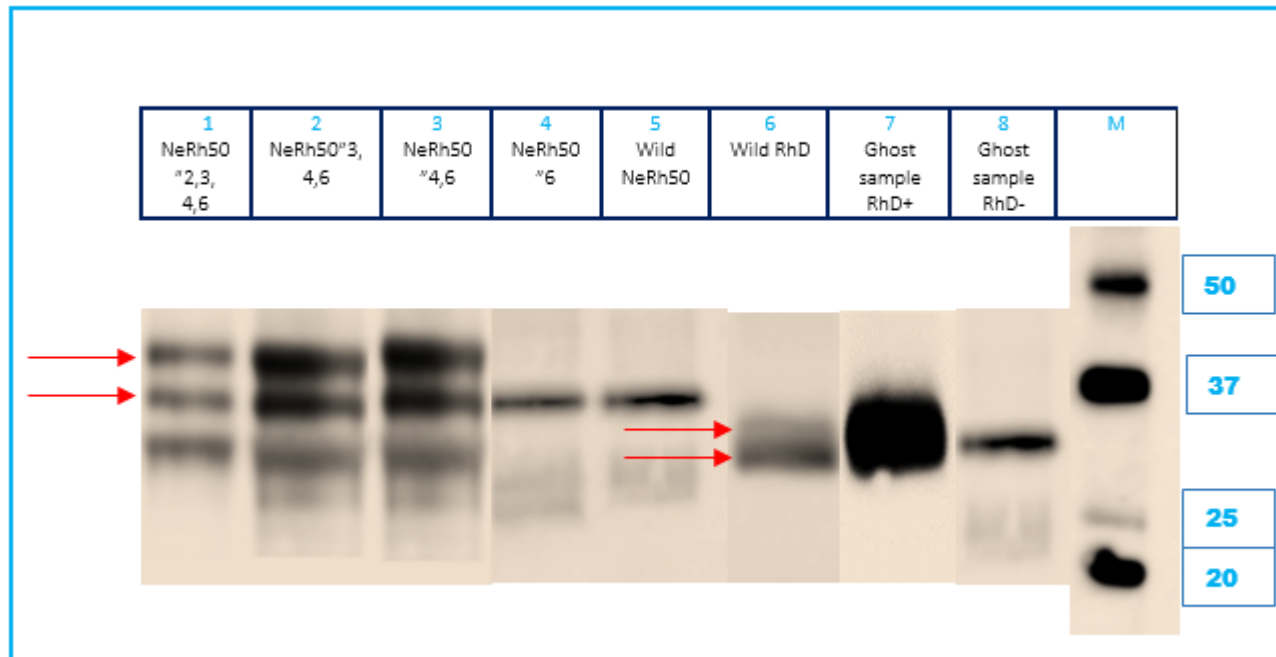


Figure 3.30. Immunostaining of RBC membranes and *E. coli* protein extractions to detect human RhD external loop 4 by using 2.2B9 antibody.

Immunostaining of RBC membranes from different Rh phenotype erythrocytes and *E. coli* protein extractions using 2.2B9. Lane (1) is representing a bacterial recombinant protein extraction of *E. coli* NeRh50^{2, 3, 4, 6} construct. Lane (2) is representing a bacterial recombinant protein extractions of *E. coli* NeRh50^{3, 4, 6} construct. Lane (3) is representing a bacterial recombinant protein extractions of *E. coli* NeRh50^{4, 6} construct. Lane (4) is representing a bacterial recombinant protein extractions of *E. coli* NeRh50⁶ construct. Lane (5) *E. coli* protein extractions for NeRh50 wild type. Lane (6) RhD30 wild type. Lane numbers (7) is representing RBC ghost membrane with Rh⁺ phenotype. Lane (8) is RBC ghost membrane with Rh⁻ phenotype. 10 μ g of RBC membrane proteins and 5 μ g of *E. coli* protein extractions were used in all samples under reducing conditions. Rh⁺ (lane 7) and the wild type human RhD (lane 6) samples show higher density bands of the human RhD external loop 4. Human RhD external loop 4 was observed around 32 kDa as a double bands, while in NeRh50^{4, 6}, NeRh50^{3, 4, 6}, and NeRh50^{2, 3, 4, 6} samples (lanes 3, 2, 1 respectively), showed a clear double bands around 37 kDa. Furthermore, NeRh50 wild type (lane 5) and NeRh50⁶ lane (4) show a similar faint band around 37 kDa. Rh⁻ lane(8) shows a one band around 32 kDa. Products were run on a 4-12 % acrylamide gel alongside a protein marker (M). Shown to the right of the figure are marker sizes in kDa. The primary antibody concentration was 1:6 and the secondary antibody was rabbit anti-mouse. Arrows are pointing to the bands of interest.

3.5.1.3 Detection of human RhD external loop 6 by 2A4 / 5G9 antibodies by western blotting

2A4 / 5G9 are two clones of a culture supernatant monoclonal antibody (see 2.1.1) and were used to detect the expression of human RhD external loop 6 in all constructs and control samples (Figure 3.31 and 3.32).

NeRh50 wild type, which was used as a negative control for both 2A4 and 5G9 antibodies, shows no obvious band. Ghost membrane sample with Rh⁻ phenotype, which was used as a second negative control, shows no band.

On the other hand, ghost membrane sample with Rh⁺ phenotype, which was used as a positive control for this antibody, shows a faint band around 32 kDa.

Moreover, all constructs; pET*NeRh50*"6 sample, pET *NeRh50*"4, 6, pET *NeRh50*"3, 4, 6, and pET *NeRh50*"2, 3, 4, 6, react similar and show a clear band around 37 kDa (Figure 3.31 and 3.32).



Figure 3.31. Immunostaining of RBC membranes and *E. coli* protein extractions to detect human RhD external loop 6 by using 2A4 antibody.

Immunostaining of RBC membranes from different Rh phenotype erythrocytes and *E. coli* protein extractions using 2A4. Lane (1) is representing a bacterial recombinant protein extraction of *E. coli* NeRh50^{2, 3, 4, 6} construct. Lane (2) is representing a bacterial recombinant protein extractions of *E. coli* NeRh50^{3, 4, 6} construct. Lane (3) is representing a bacterial recombinant protein extractions of *E. coli* NeRh50^{4, 6} construct. Lane (4) is representing a bacterial recombinant protein extractions of *E. coli* NeRh50⁶ construct. Lane (5) *E. coli* protein extractions for NeRh50 wild type. Lanes (6), (7) is representing RBC ghost membrane with Rh⁺ phenotype and RBC ghost membrane with Rh⁻ phenotype respectively. 10 μ g of RBC membrane proteins and 5 μ g of *E. coli* protein extractions were used in all samples under reducing conditions. Rh⁺ lane (6) shows a faint band for the human RhD external loop 6. Human RhD external loop 6 was observed around 32 kDa as single band, while in NeRh50⁶, NeRh50^{4, 6}, NeRh50^{3, 4, 6}, and NeRh50^{2, 3, 4, 6} samples (lanes 4, 3, 2, 1 respectively), showed a clear band around 37 kDa. Furthermore, NeRh50 wild type lane (5) showed no band. Rh⁻ lane (7) shows a no band. Products were run on a 4-12 % acrylamide gel alongside a protein marker (M). Shown to the right of the figure are marker sizes in kDa. The primary antibody concentration was 1:10 and the secondary antibody was rabbit anti-mouse. Arrows are pointing to the interested bands

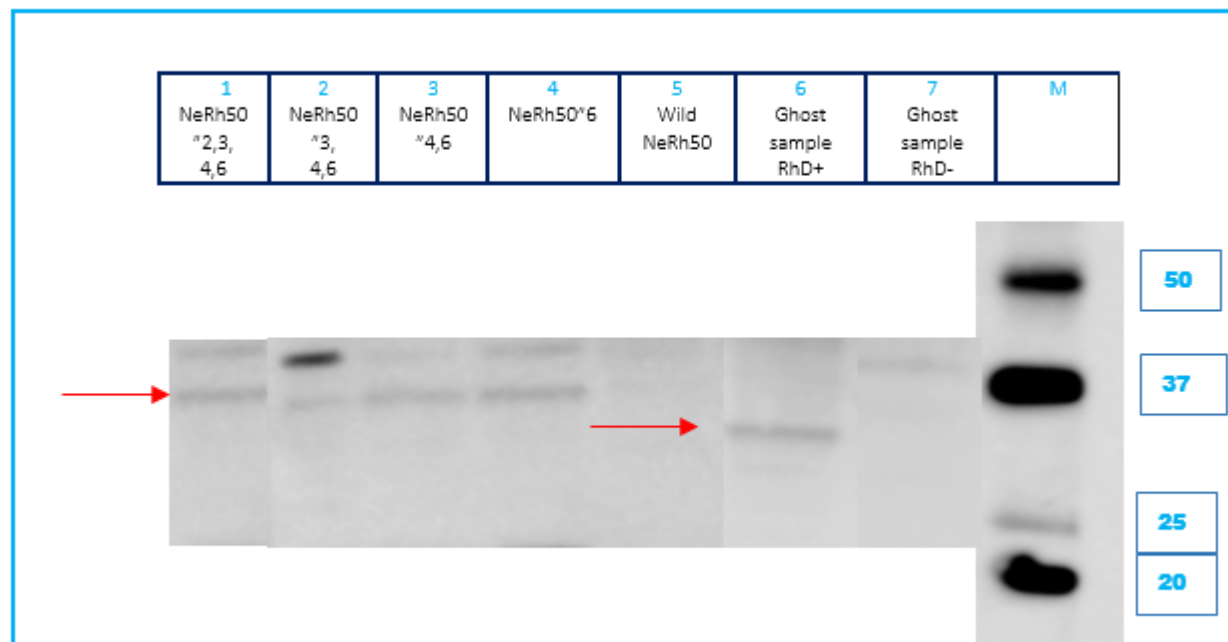


Figure 3.32. Immunostaining of RBC membranes and *E. coli* protein extractions to detect human RhD external loop 6 by using 5G9 antibody.

Immunostaining of RBC membranes from different Rh phenotype erythrocytes and *E. coli* protein extractions using 5G9. Lane (1) is representing a bacterial recombinant protein extractions of *E. coli* NeRh50^{2, 3, 4, 6} construct. Lane (2) is representing a bacterial recombinant protein extractions of *E. coli* NeRh50^{3, 4, 6} construct. Lane (3) is representing a bacterial recombinant protein extractions of *E. coli* NeRh50^{4, 6} construct. Lane (4) is representing a bacterial recombinant protein extractions of *E. coli* NeRh50⁶ construct. Lane (5) *E. coli* protein extractions for NeRh50 wild type. Lanes (6), (7) is representing RBC ghost membrane with Rh⁺ phenotype and RBC ghost membrane with Rh⁻ phenotype respectively. 10 μ g of RBC membrane proteins and 5 μ g of *E. coli* protein extractions were used in all samples under reducing conditions. Rh⁺ lane (6) shows a faint band for the human RhD external loop 6. Human RhD external loop 6 was observed around 32 kDa as a single band, while in NeRh50⁶, NeRh50^{4, 6}, NeRh50^{3, 4, 6}, and NeRh50^{2, 3, 4, 6} samples (lane 4, 3, 2, 1 respectively), showed a faint band around 37 kDa. Furthermore, NeRh50 wild type lane (5) shows no band. Rh⁻ lane (7) shows no band. Products were run on a 4-12 % acrylamide gel alongside a protein marker (M). Shown to the right of the figure are marker sizes in kDa. The primary antibody concentration was 1:10 and the secondary antibody was rabbit anti-mouse. Arrows are pointing to the interested bands

3.5.1.4 Detection of human RhD external loops 4, 6 by BIRMA-D6 antibody by western blotting:

BIRMA-D6 is a purified monoclonal antibody (see 2.1.1), which was used to detect the expression of human RhD external loops 4, 6 in all constructs and control samples (Figure 3.33).

NeRh50 wild type, which was used as a negative control for BIRMA-D6, it shows no band. Ghost membrane sample with Rh⁻ phenotype, which was used as a second negative control for BIRMA-D6, shows no obvious band.

On the other hand, ghost membrane sample with Rh⁺ phenotype, which was used as a positive control to this antibody, shows a clear band around 32 kDa.

Furthermore, all constructs; pET *NeRh50*"6, pET *NeRh50*"4, 6, pET *NeRh50*"3, 4, 6, and pET *NeRh50*"2, 3, 4, 6, react similar and show a band around 37 kDa (Figure 3.33).

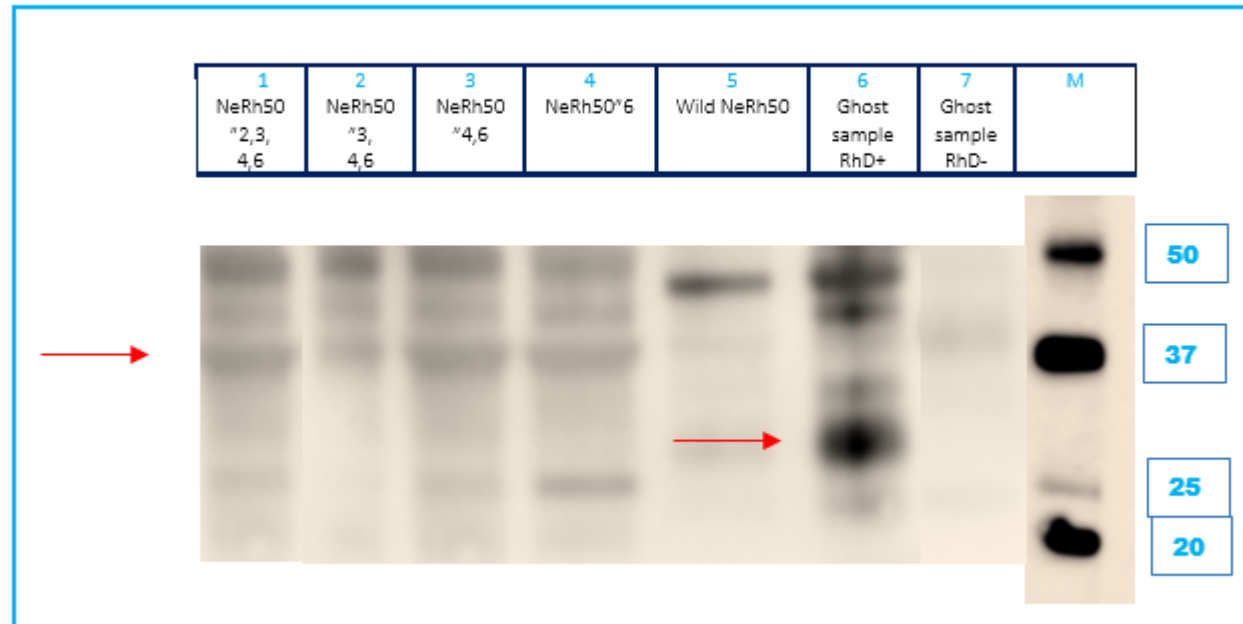


Figure 3.33. Immunostaining of RBC membranes and *E. coli* protein extractions to detect human RhD external loops 4, 6 by using BIRMA-D6 anti-body.

Immunostaining of RBC membranes from different Rh phenotypes erythrocytes and *E. coli* protein extractions using BIRMA-D6. Lane (1) is representing a bacterial recombinant protein extractions of *E. coli* NeRh50^{2, 3, 4, 6} construct. Lane (2) is representing a bacterial recombinant protein extractions of *E. coli* NeRh50^{3, 4, 6} construct. Lane (3) is representing a bacterial recombinant protein extractions of *E. coli* NeRh50^{4, 6} construct. Lane (4) is representing a bacterial recombinant protein extractions of *E. coli* NeRh50⁶ construct. Lane (5) *E. coli* protein extractions for NeRh50 wild type. Lanes (6), (7) is representing RBC ghost membrane with Rh⁺ phenotype and RBC ghost membrane with Rh⁻ phenotype respectively. 10 μ g of RBC membrane proteins and 5 μ g of *E. coli* protein extractions were used in all samples under reducing conditions. Rh⁺ lane (6) shows a clear band for the human RhD external loop 6. Human RhD external loops 4, 6 were observed around 32 kDa as single band, while in NeRh50⁶, NeRh50^{4, 6}, NeRh50^{3, 4, 6}, and NeRh50^{2, 3, 4, 6} samples (lanes 4, 3, 2, 1 respectively), showed a clear band around 37 kDa. Furthermore, NeRh50 wild type lane (5) showed no band at the same size. Rh⁻ lane (7) shows a no band. Products were run on a 4-12 % acrylamide gel alongside a protein marker (M). Shown to the right of the figure are marker sizes in kDa. The primary antibody concentration was 1:2500 and the secondary antibody was goat anti-human. Arrows are pointing to the band of interest.

3.5.1.5 Detection of human RhD external loops 3, 4, 6 by MS-26 antibody by western blotting:

MS-26 is a culture supernatant monoclonal antibody (see 2.1.1) was used to detect the expression of human RhD external loops 3, 4, 6 in all constructs and control samples (Figure 3.34).

It was raised against a synthetic peptide corresponding to EpD23 of the RhD protein (Scott,1996).

NeRh50 wild type, which was used as a negative control for MS-26, shows no band. Ghost membrane sample with Rh⁻ phenotype, which was used as a second negative control to MS-26, shows no obvious band, reflecting that this MAb may recognise a conformation- dependent antibody, not detectable by western blotting. On the other hand, ghost membrane sample with Rh⁺ phenotype, and all constructs; pET*NeRh50*"6, pET *NeRh50*"4, 6, pET *NeRh50*"3, 4, 6, and pET *NeRh50*"2, 3, 4, 6, show no band at any size (Figure 3.34).



Figure 3.34. Immunostaining of RBC membranes and *E. coli* protein extractions to detect human RhD external loops 3, 4, 6 by using MS-26 antibody.

Immunostaining of RBC membranes from different Rh phenotypes erythrocytes and *E. coli* protein extractions using MS-26. Lane (1) is representing a bacterial recombinant protein extractions of *E. coli* NeRh50"2, 3, 4, 6 construct. Lane (2) is representing a bacterial recombinant protein extractions of *E. coli* NeRh50"3, 4, 6 construct. Lane (3) is representing a bacterial recombinant protein extractions of *E. coli* NeRh50"4, 6 construct. Lane (4) is representing a bacterial recombinant protein extractions of *E. coli* NeRh50"6 construct. Lane (5) *E. coli* protein extractions for NeRh50 wild type. Lanes (6), (7) is representing RBC ghost membrane with Rh⁺ phenotype and RBC ghost membrane with Rh⁻ phenotype respectively .

Products were run on a 4-12 % acrylamide gel percentage alongside a protein marker (M). Shown to the right of the figure are marker sizes in kDa. 10 μ g of RBC membrane proteins and 5 μ g of *E. coli* protein extractions were used in all samples under reducing conditions. All controls and constructs were shown no obvious bands. The primary antibody concentration was 1:2500 and the secondary antibody was Rabbit anti-mouse.

3.5.1.6 Detection of NeRh50 external loop 1 and C-terminus by 1918 antibody by western blotting:

1918 is a purified polyclonal antibody (see 2.1.1) was used to detect the expression of *NeRh50* external loop 1 and C-terminus in all constructs and control samples. It was raised against a synthetic peptide corresponding to the first external loop and C-terminus domains of the *NeRh50* protein (Figure 3.35).

Ghost membrane samples with Rh⁻ and Rh⁺ phenotypes, which were used as a negative control to 1918 antibody, shows no obvious band.

By contrast, *NeRh50* wild type, which used as a positive control to this antibody, shows a clear band around 37 kDa.

Furthermore, all constructs; pET*NeRh50*"6, pET *NeRh50*"4, 6, pET *NeRh50*"3, 4, 6, and pET *NeRh50*"2, 3, 4, 6, react similar to the positive control *NeRh50* wild type and show a band around 37 kDa (Figure 3.35).

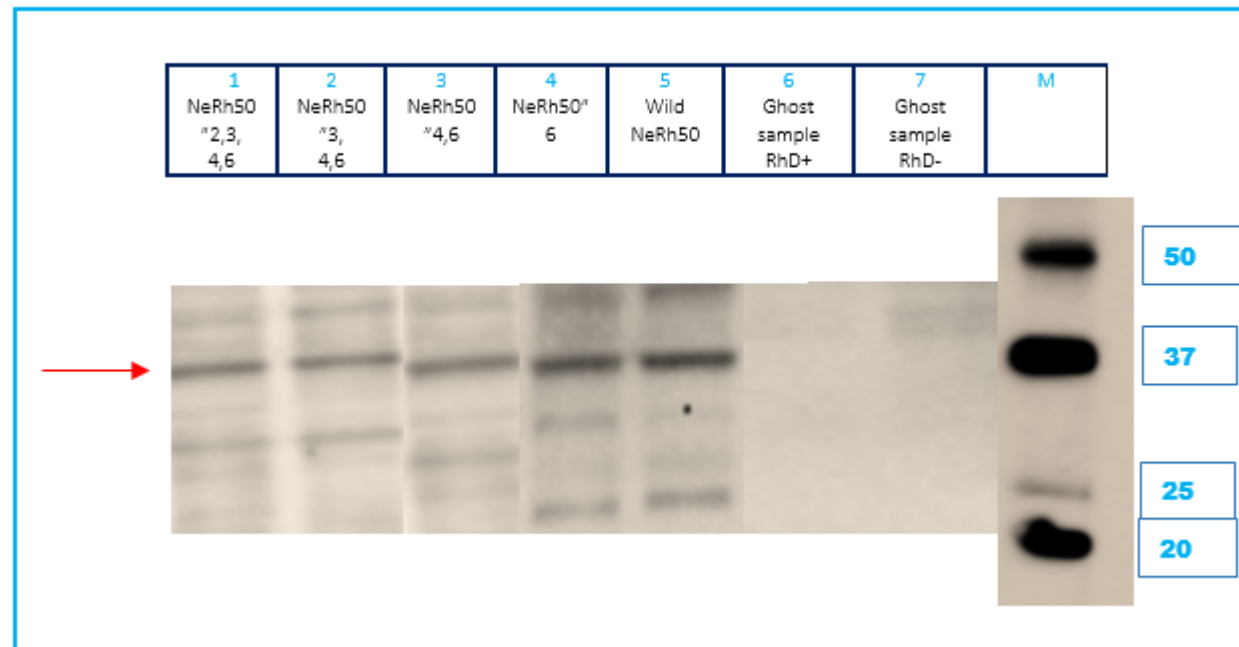


Figure 3.35. Immunostaining of RBC membranes and *E. coli* protein extractions to detect NeRh50 external loop 1 and C-terminus by using 1918 antibody.

Immunostaining of RBC membranes from different Rh phenotypes erythrocytes and *E. coli* protein extractions using 1918.

Lane (1) is representing a bacterial recombinant protein extractions of *E. coli* NeRh50^{"2, 3, 4, 6} construct. Lane (2) is representing a bacterial recombinant protein extractions of *E. coli* NeRh50^{"3, 4, 6} construct. Lane (3) is representing a bacterial recombinant protein extractions of *E. coli* NeRh50^{"4, 6} construct. Lane (4) is representing a bacterial recombinant protein extractions of *E. coli* NeRh50^{"6} construct. Lane (5) *E. coli* protein extractions for NeRh50 wild type. Lanes (6), (7) is representing RBC ghost membrane with Rh⁺ phenotype and RBC ghost membrane with Rh⁻ phenotype respectively. 10 μ g of RBC membrane proteins and 5 μ g of *E. coli* protein extractions were used in all samples under reducing conditions. Rh⁺, Rh⁻ lanes (6, 7) shows no band. NeRh50 external loop 1 and C-terminus were observed around 37 kDa as single band as appear in; NeRh50 wild type NeRh50^{"6}, NeRh50^{"4, 6}, NeRh50^{"3, 4, 6}, and NeRh50^{"2, 3, 4, 6} samples (lanes 5, 4, 3, 2, 1 respectively), showed a clear band around 37 kDa. Products were run on a 4-12 % acrylamide gel percentage alongside a protein marker (M). Shown to the right of the figure are marker sizes in kDa. The primary antibody concentration was 1:2500 and the secondary antibody was swine anti-rabbit. Arrows are pointing to the band of interest.

3.5.1.7 Detection of NeRh50 external loop 1 by 1917 antibody by western blotting:

1917 is a purified monoclonal antibody (see 2.1.1) was used to detect the expression of *NeRh50* external loop 1 in all constructs and control samples. It was raised against a synthetic peptide corresponding to the first external domains of the *NeRh50* protein (Figure 3.36).

Ghost membrane samples with Rh⁻ and Rh⁺ phenotypes, which were used as a negative control to 1918 antibody, shows no obvious band.

By contrast, *NeRh50* wild type, which used as a positive control for this antibody, shows a clear band around 37 kDa. Furthermore, all constructs; pET*NeRh50*"6, pET *NeRh50*"4, 6, pET *NeRh50*"3, 4, 6, and pET *NeRh50*"2, 3, 4, 6, react similar to the positive control *NeRh50* wild type and show a band around 37 kDa (Figure 3.36). 1917 blots with wild type NeRh50/ NeRh50"6/ NeRh50"4, 6/ NeRh50"3, 4, 6 / NeRh50"2, 3, 4, 6, samples, which showed a clear band around 37 kDa compared to Rh⁺, wild type human RhD samples and Rh⁻ sample, which shows nothing around the same size.



Figure 3.36. Immunostaining of RBC membranes and *E. coli* protein extraction using 1917.

Immunostaining of RBC membranes from different Rh phenotypes erythrocytes and *E. coli* protein extractions using 1917. Lane numbers (1-4) are *E. coli* protein extraction for constructs (NeRh50"2, 3, 4, 6 NeRh50"3, 4, 6/ NeRh50"4, 6 /NeRh50"6), lane(5) *E. coli* protein extraction for wild type NeRh50, which shows no band. Lane numbers (6, 7) are representing RBC membranes for Rh⁺, Rh⁻ respectively. 10 μ g of RBC membrane proteins and 5 μ g of *E. coli* protein extractions were used in all samples under reducing conditions. Rh⁺, Rh⁻ lanes (6, 7) shows no band. *NeRh50* external loop 1 and C-terminus were observed around 37 kDa as single band as appear in; *NeRh50* wild type *NeRh50*"6, *NeRh50*"4, 6, *NeRh50*"3, 4, 6, and *NeRh50*"2, 3, 4, 6 samples (lanes 5, 4, 3, 2, 1 respectively), showed a clear band around 37 kDa. Products were run on an acrylamide gel 4-12% alongside a protein marker (M). Shown to the right of the figure are maker sizes in kDa. The primary antibody concentration was 1:2500 and the secondary antibody was rabbit anti-mouse. Arrows are pointing to the band of interest.

3.5.2 Flow cytometry Results

Flow cytometry was completed using a total of 11 different monoclonal anti-D and anti-*NeRh50* antibodies shown in Table 2.1 and using the methods outlined (See 2.2.2.4.9) to analyse the different expression patterns of D epitopes and *NeRh50* loop 1, C-terminus. These antibodies were used in conjunction with a FITC assay and a FACS ARIA II Flow cytometer/Cell Sorter device were used. Four different constructs; pET*NeRh50*"6, pET *NeRh50*"4, 6, pET *NeRh50*"3, 4, 6, pET *NeRh50*"2, 3, 4, 6, RhD30 wild type and *NeRh50* wild type clones were transformed into BL21 (DE3) *E. coli* competent cells (see 2.2.2.4.4). Two types of negative control were used: non stained samples and samples used with secondary FITC-conjugated antibodies (See 2.2.2.4.9.2). Moreover, negative controls were used to assess the background fluorescence by comparing these reads to the reads from the samples treated with primary and secondary antibodies. Mean fluorescence intensity was used to measure the antibody binding. Table 3.4 shows average mean fluorescence intensity (MFI) values of screening the four different constructs; pET*NeRh50*"6, pET *NeRh50*"4, 6, pET *NeRh50*"3, 4, 6, pET *NeRh50*"2, 3, 4, 6, RhD30 wild type, and *NeRh50* wild type. Distinct patterns of reactivity were observed, ranging from the complete absence of D epitope expression in the case of *NeRh50* wild type sample to weak and strong reaction of constructs and RhD30 wild type to anti-D (Figure 3.39, 3.40, 3.41).

Human RhD30 wild type showed the highest expression when using the BRAD-3 monoclonal anti-D while shows low expression when it reacts with 1918, which is anti-*NeRh50*. In contrast, *NeRh50* wild type sample showed the highest read with 1918 and lowest expression with HAM-A, which is a monoclonal anti-D EpD 11 (Scott, 1996).

Anti-*NeRh50* 1918 has the highest expression with *NeRh50*"4, 6 construct while it has the lowest expression with *NeRh50*"2, 3, 4, 6 construct.

Anti-D 2.2B9 shows the highest expression with *NeRh50*"2, 3, 4, 6 construct, however, it shows the lowest read with *NeRh50*"6. *NeRh50*"3, 4, 6 construct shows the lowest expression with all anti-D antibodies, while it shows moderate expression with anti-*NeRh50* 1918.

Figures 3.37 and 3.38 illustrate the mean Fluorescence intensity (MFI) values by three dimensional chart for anti-D antibodies against constructs of a recombinant protein *NeRh50* human RhD external loops, RhD30 wild type and *NeRh50* wild type. Moreover, comparing their reads to un-stained samples and samples treated with secondary antibodies only. In these charts the variation of the reaction of anti-D and anti-*NeRh50* against control samples and constructs appear as different bar length.

Antibodies	Anti-EpD1-37	RhD30	NeRh50	NeRh50 6	NeRh50 4, 6	NeRh50 3, 4, 6	NeRh50 2, 3, 4, 6
2.2B9		155	25	9	504	129	773
MS-26	EpD 23	627	67	224	353	103	1049
2A4		161	39	129	343	26	281
5G9		51	13	90	705	70	157
BIRMA-D6	EpD 23	1358	163	1683	613	388	1052
BIRMA-D10	EpD 8	569	15	580	81	73	873
BRAD-3	EpD 13	3668	15	2338	1552	935	1378
RUM-1	EpD 6/7	253	15	870	1554	94	696
HAM-A	EpD 11/5	476	4	621	879	199	845
MAD-2	EpD 19	547	10	317	697	35	860
1918		60	428	478	933	903	463

Table 3.4. Flow Cytometric Analysis of wild type RhD, wild type *NeRh50* and 4 different constructs using Monoclonal Anti-D and Anti-*NeRh50*.

Table shows the average of triplicate mean fluorescence intensity (MFI) values of NeRh50/human RhD external loop constructs, RhD30 and NeRh50 wild type. The epitope specificity of each anti-D is shown using the 1-37 model (Scott, 1996).

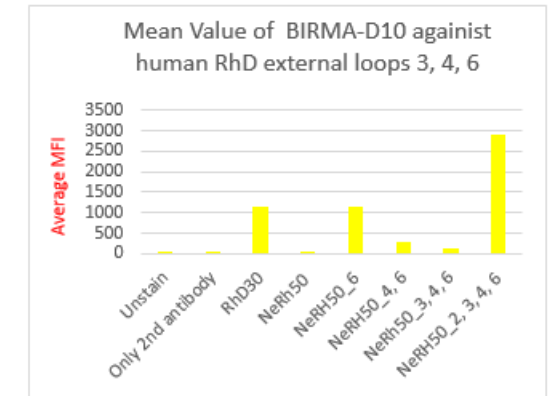
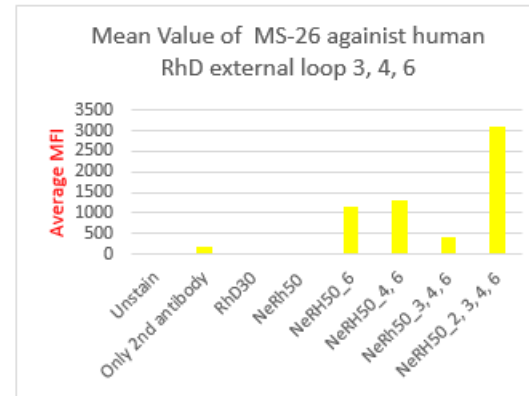
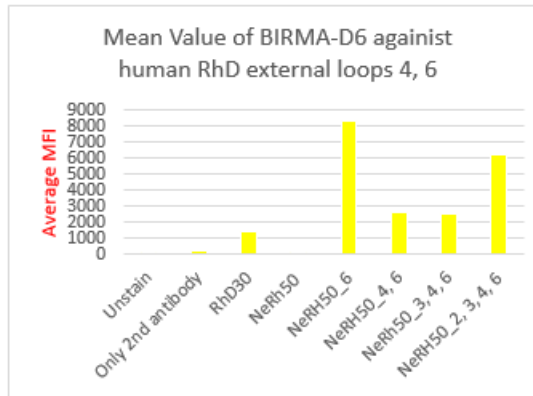
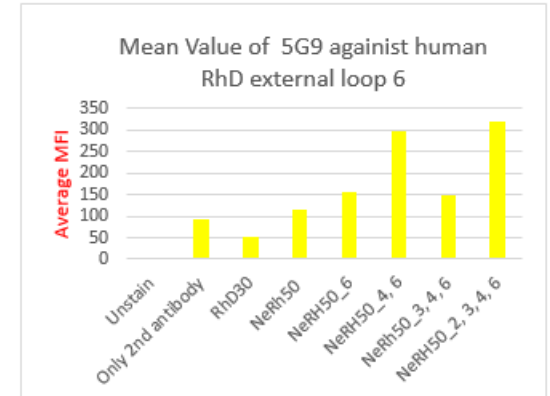
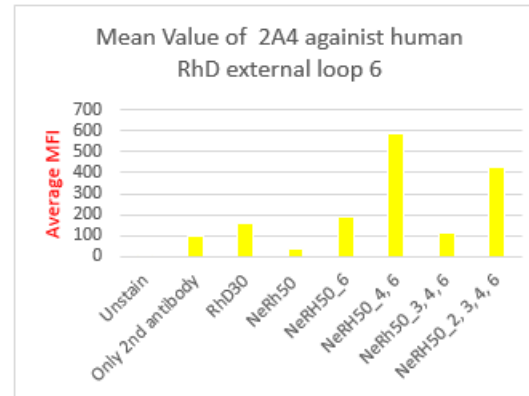
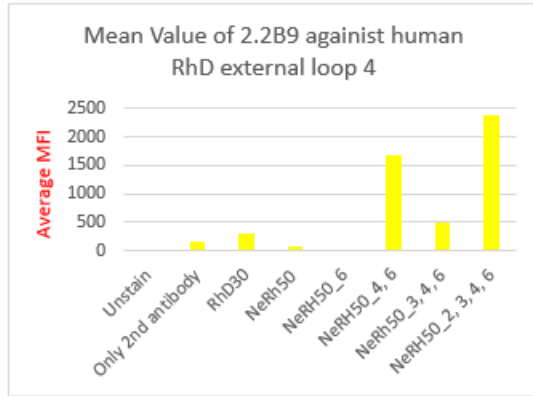




Figure 3.37. MFI values of several Anti-D against various NeRh50/human RhD external loop constructs

3-D clustered column chart shows the average reads of mean FITC Fluorescence value for different Anti-D (2.2B9, MS-26, 2A4, BIRMA-D6, 5G9, BRAD-3, HAM-A, RUM-1, BIRMA-D10, and MAD-2) against un stain *E. coli* cells work as a negative control. *E. coli* cells treated with only secondary-FITC conjugated antibody as a negative control. *E. coli* cell contain RhD30 wild type as a positive control. *E. coli* cell contain NeRh50 wild type as a negative control. compared to all Recombinant constructs *E. coli* cell contain; *NeRh50*₆, *NeRh50*_{4, 6}, *NeRh50*_{3, 4, 6}, *NeRh50*_{2, 3, 4, 6}.

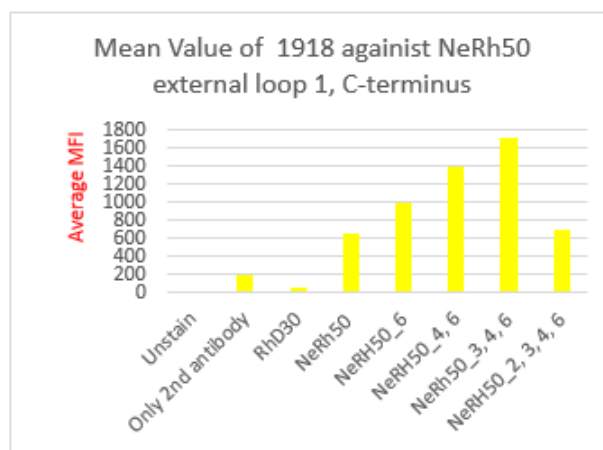


Figure 3.38. MFI values of Anti-*NeRh50* (1918) against various NeRh50/human RhD external loop constructs

3-D clustered column chart shows the average reads of mean FITC Fluorescence value for of Anti-*NeRh50* (1918) against un stain *E. coli* cells work as a negative control. *E. coli* cells treated with only secondary-FITC conjugated antibody as a negative control. *E. coli* cell contain RhD30 wild type as a negative control. *E. coli* cell contain NeRh50 wild type as a positive control. compared to all Recombinant constructs *E. coli* cell contain *NeRh50*"6, *NeRh50*"4, 6, *NeRh50*"3, 4, 6, *NeRh50*"2, 3, 4, 6.

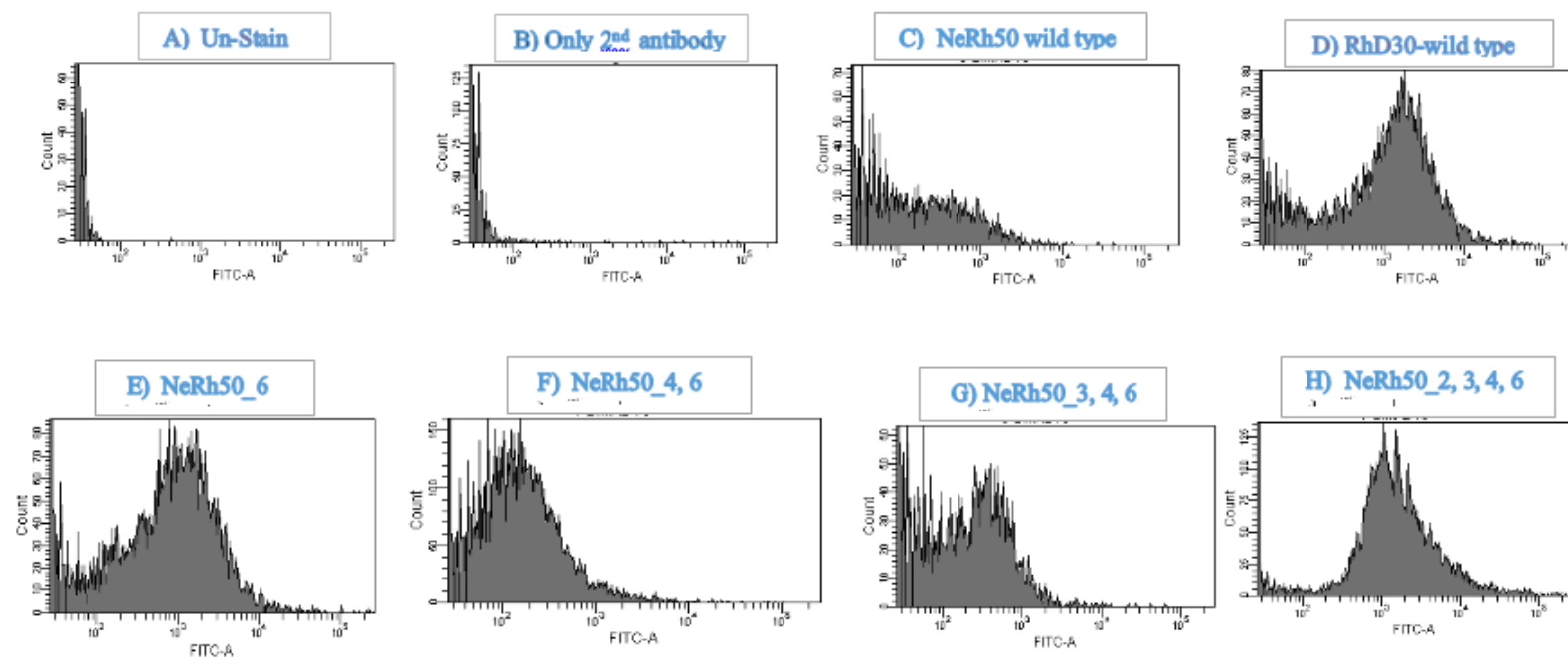


Figure 3.39. FITC fluorescence histogram of *NeRh50* human RhD external loops with Anti-D BIRMA-D10 (EpD8). FITC fluorescence histogram of *NeRh50* human RhD external loops in *E. coli* using Anti-D BIRMA-D10. A) Un-stain *E. coli* cells work as a negative control. B) *E. coli* cells work as a negative control by using only secondary antibody. C) *E. coli* contain *NeRh50* wild type work as a negative control with Anti-D BIRMA-D10. D) *E. coli* contains RhD30 wild type work as a positive control with Anti-D BIRMA-D10. E) *E. coli* contain *NeRh50* 6. F) *E. coli* contains *NeRh50* 4, 6. G) *E. coli* contains *NeRh50* 3, 4, 6. H) *E. coli* contains *NeRh50* 2, 3, 4, 6.

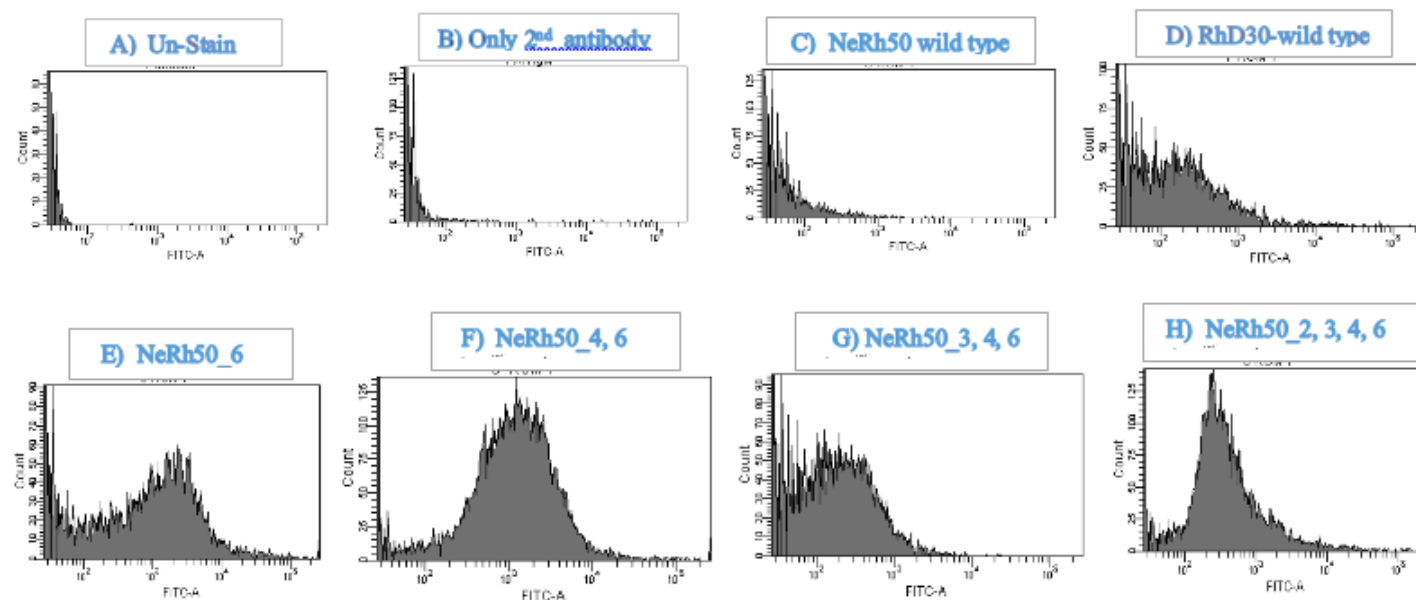


Figure 3.40. FITC fluorescence histogram of *NeRh50* human RhD external loops with Anti-D RUM-1 (EpD 6/7)

FITC fluorescence histogram of *NeRh50*/human RhD external loops in *E. coli* using Anti-D RUM-1. A) Un-stain *E. coli* cells work as a negative control. B) *E. coli* cells work as a negative control by using only secondary antibody. C) *E. coli* contains *NeRh50* wild type work as a negative control with Anti-D RUM-1. D) *E. coli* contain RhD30 wild type work as a positive control with Anti-D RUM-1. E) *E. coli* contains *NeRh50*⁶. F) *E. coli* contains *NeRh50*^{4, 6}. G) *E. coli* contains *NeRh50*^{3, 4, 6}. H) *E. coli* contains *NeRh50*^{2, 3, 4, 6}.

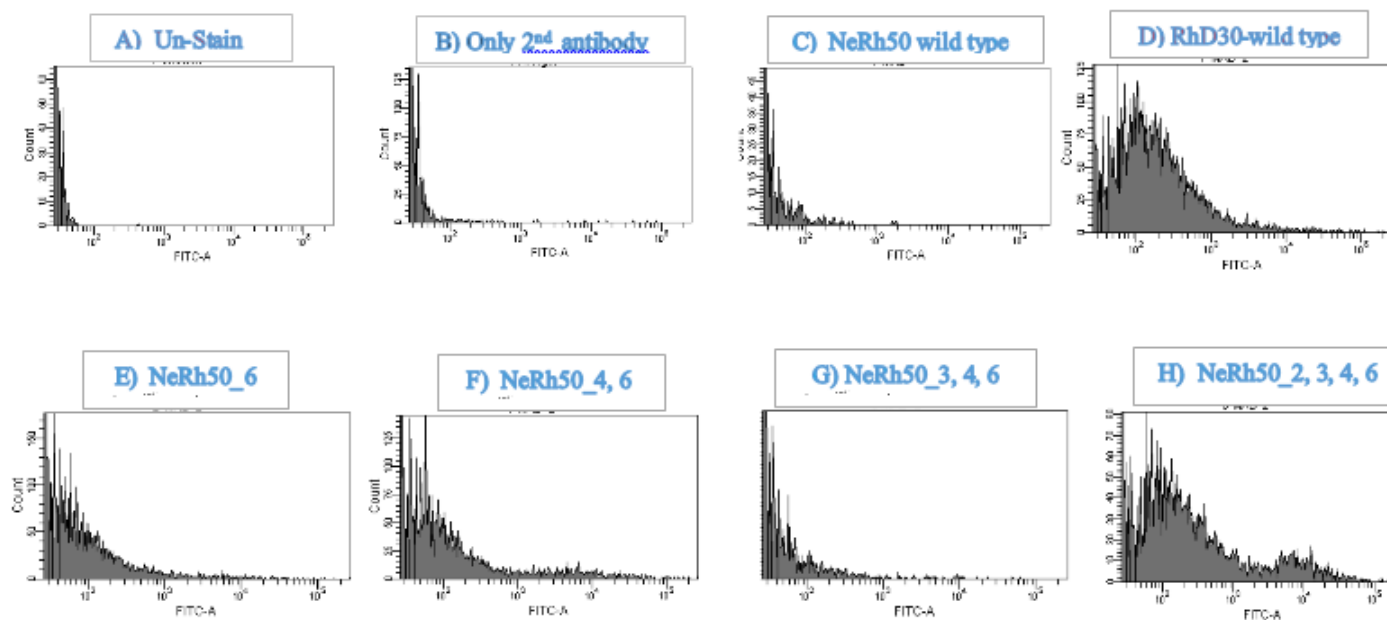


Figure 3.41. FITC fluorescence histogram of *NeRh50*_human RhD external loops with Anti-D MAD-2 (EpD 19)

FITC fluorescence histogram of *NeRh50*_human RhD external loops in *E. coli* using Anti-D MAD-2. A) Un-stain *E. coli* cells work as a negative control. B) *E. coli* cells work as a negative control by using only secondary antibody. C) *E. coli* contain *NeRh50* wild type work as a negative control with Anti-D MAD-2. D) *E. coli* contain RhD30 wild type work as a positive control with Anti-D MAD-2. E) *E. coli* contain *NeRh50*"6. F) *E. coli* contains *NeRh50*"4, 6. G) *E. coli* contains *NeRh50*"3, 4, 6. H) *E. coli* contains *NeRh50*"2, 3, 4, 6.

3.5.3 Transmission Electron Microscopy Results

Immuno-gold staining was also used to confirm the presence of the recombinant *RHD* gene in different constructs; pET *NeRh50*"6, pET *NeRh50*"4, 6, pET *NeRh50*"3, 4, 6, and pET *NeRh50*"2, 3, 4, 6, by detecting the expression of RhD on the *E. coli* membrane by using anti-D 2.2B9 against loop 4 and 2A4 against loop 6. Anti-*NeRh50* 1918 against loop 1 and C-terminus was used to detect wild type *NeRh50* protein. Furthermore, an anti *E. coli* antibody was used as a positive control. TEM was used to visualize and image the 10 nm (colloidal gold) trace, which was conjugated to the secondary antibodies.

There was a significant change in the level of expression between the antibodies; 2.2B9, which should detect human RhD external loop 4 showed no binding in *NeRh50*"6 construct similar to the negative control image. While, 2.2B9 works with other constructs pET *NeRh50*"4, 6, pET *NeRh50*"3, 4, 6, and pET *NeRh50*"2, 3, 4, 6 indicate the presence of human RhD external loop 4 in these constructs. pET *NeRh50*"2, 3, 4, 6 construct shows the highest expression level for human RhD external loop 4, which appear as black dots of gold trace (Figure 3.42).

2A4, which is able to detect human RhD external loop 6 showed a high expression level in all constructs. Moreover, the *E. coli* protein showed high expression with anti- *E. coli* antibody in pET *NeRh50* wild type sample, which was used as a positive control (Figure 3.43).

Figure 3.44, all constructs reacted at a moderate level with anti-*NeRh50* 1918, which detects *NeRh50* external loop 1 and C-terminus. *NeRh50"6* shows the lowest level of the expression for *NeRh50* external loop1 and C-terminus compared to other constructs.

2.2B9 (Figure 3.42) shows low levels of expression for human RhD external loop 4 compared to 2A4 for human RhD external loop 6 (Figure 3.43), which consistently demonstrated high levels of expression across all constructs. 1918 demonstrated a lower level of binding than 2A4 but more than 2.2B9 (Figure 3.44).

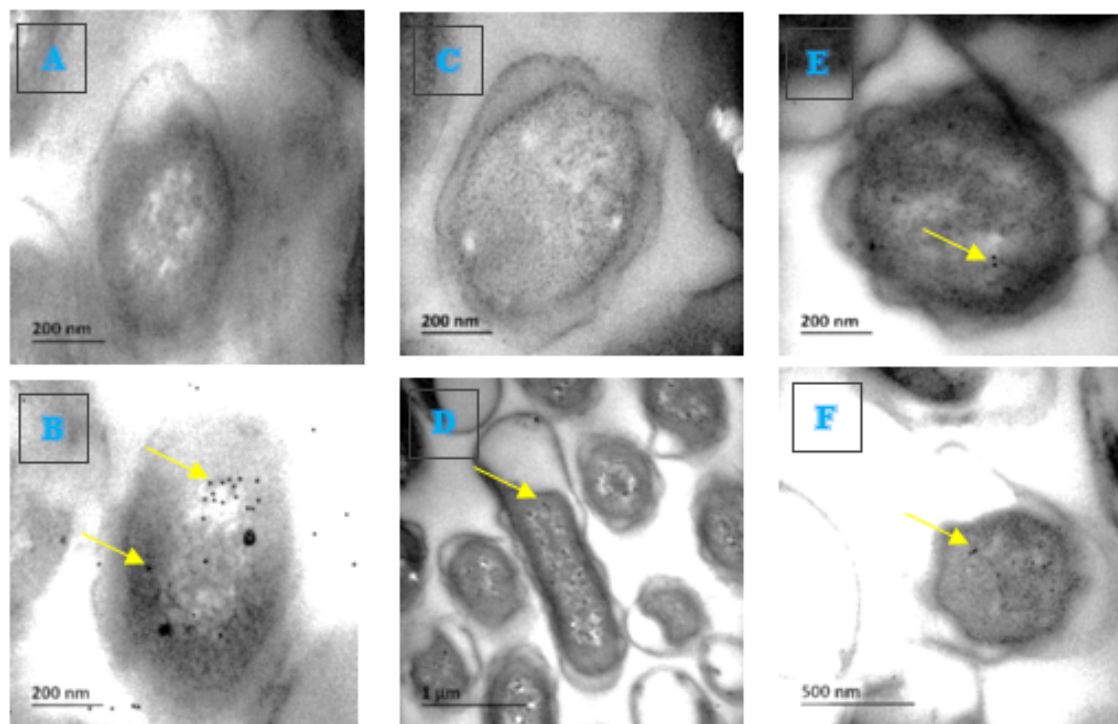


Figure 3.42. Transmission electron micrograph for a thin layer of *E. coli* contains recombinant *NeRh50*"RhD by using anti-D loop 4 2.2B9.

TEM images of *E. coli* cells contain different constructs of *NeRh50*"human RhD external loops. Black dots are representing gold trace, which conjugated to the secondary antibody to show the protein presence in the *E. coli* membrane. A) *E. coli* cells work as a negative control by using only secondary antibody, which show no dots. B) *E. coli* contain *NeRh50* wild type with Anti.*E. coli*, which work as a positive control, it shows many black dots with different intensity. C) *E. coli* contains *NeRh50*"6, which shows no gold trace. D) *E. coli* contain *NeRh50*"4, 6, which shows gold trace. E) *E. coli* contain *NeRh50*"3, 4, 6. F) *E. coli* contain *NeRh50*"2, 3, 4, 6, which shows gold trace.

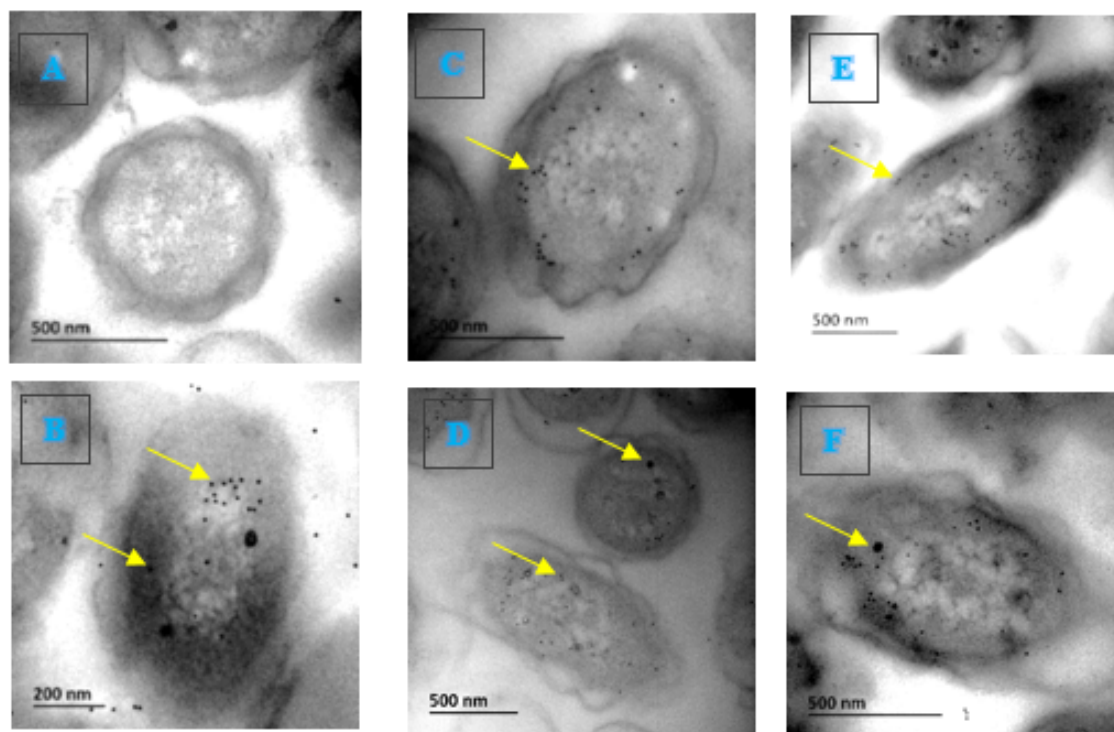


Figure 3.43. Transmission electron micrograph for a thin layer of *E. coli* contains recombinant *NeRh50*_RhD by using anti-D loop 6 2A4. TEM images of *E. coli* cells contain different constructs of *NeRh50* human RhD external loops. Black dots are representing gold trace, which conjugated to the secondary antibody to show the protein presence in the *E. coli* membrane. A) *E. coli* cells work as a negative control by using only secondary antibody, which show no dots. B) *E. coli* contain *NeRh50* wild type with Anti.*E. coli* work as a positive control, which shows many black dots with different intensity. C) *E. coli* contain *NeRh50*"6, which shows gold trace. D) *E. coli* contain *NeRh50*"4, 6, which shows gold trace. E) *E. coli* contain *NeRh50*"3, 4, 6. F) *E. coli* contain *NeRh50*"2, 3, 4, 6, which shows gold trace.

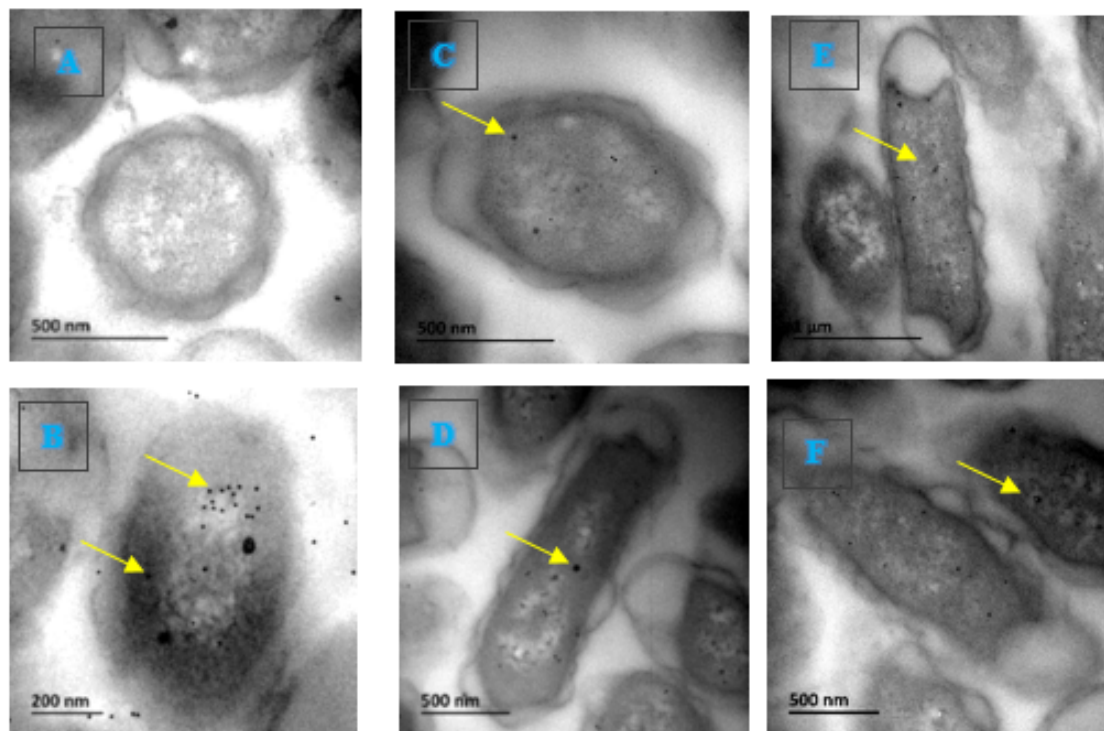


Figure 3.44. Transmission electron micrograph for a thin layer of *E. coli* contains recombinant *NeRh50*"RhD by using anti-*NeRh50* loop 1 and C-terminus 1918.

TEM images of *E. coli* cells contain different constructs of *NeRh50*"human RhD external loops. Black dots are representing gold trace, which conjugated to the secondary antibody to show the protein presence in the *E. coli* membrane. A) *E. coli* cells work as a negative control by using only secondary antibody, which show no dots. B) *E. coli* contain *NeRh50* wild type with Anti-*E. coli* work as a positive control, which shows many black dots with different intensity. C) *E. coli* contain *NeRh50*"6, which shows gold trace. D) *E. coli* contain *NeRh50*"4, 6, which shows gold trace. E) *E. coli* contain *NeRh50*"3, 4, 6. F) *E. coli* contain *NeRh50*"2, 3, 4, 6, which shows gold trace.

3.6 Discussion

3.6.1 Generating constructs

This chapter starts with describing the generation of different constructs based on the *NeRh50* wild type sequence as a backbone and adding the human RhD external loops to create; *NeRh50*"6, *NeRh50*"4, 6, *NeRh50*"3, 4, 6, and *NeRh50*"2, 3, 4, 6.

The *E. coli* codon usage table was used to reverse translate the human RhD external loops to increase the chance of production of the recombinant proteins in *E. coli* and expression in the *E. coli* membrane. The codon usage in humans differs to the codon usage in *E. coli* and if human codon usage had been used it ran the risk of the recombinant proteins not being produced.

Due to the lack of vector map for the pESV2 expression vector provided by Prof Merrick's lab, NGS was used to sequence the vector and plasmid constructs. The results showed that the pESV2 vector was not a pESV2 vector and hence the expression vector for the constructs was changed to be pET-9a.

Following enzymatic digestion with *Bam*HI and *Xba*I for the recombinant constructs pET-*NeRh50*"human RhD external loops (Figure 3.21), the band sizes of the inserts were correct and hence the correct constructs have been cloned; plasmid (pET expression vector) with a band size around 3 kb and the constructs in each plasmid; *NeRh50*"6 1766 bp, *NeRh50*"4, 6 1275

bp, *NeRh50*"3, 4, 6 1275 bp, and *NeRh50*"2, 3, 4, 6 1275 bp. Moreover, Sanger sequencing data for the same colonies confirmed that the vector is, pET expression vector as expected by amplifying areas around the insertion; F1 forward primer (Figure 3.23), R1 reverse primer (Figure 3.24) to cover the region before the hybrid constructs and the alignment shows similarity between all constructs. Furthermore, F7 forward primer (Figure 3.25) and R7 reverse primer (Figure 3.26) were cover the region in the pET vector after the hybrid sequence. Although other primer sets shown in Table 2.4 were not able to amplify the interested region within the hybrid recombinant sequence (Appendix C-I).

This project was based on the successful result of SDM work by (Liu *et al.*, 1999 a, b). They were able to express *RHCE-RHD* recombinant; the construct was based on *RHCE* as backbone and replacing its third, fourth and sixth external loop with the human *RHD* corresponding external loops. D epitopes successfully resulted from utilising the human RHCE backbone with the addition of human RHD external loops 3, 4 and 6. So, a simple change in the backbone lead to produce D-epitopes, which may prove that human *RHD* external predicted domains are responsible for the production of D-epitopes. As described above synthetic construction of *NeRh50*-human RhD external loops; pET*NeRh50*"6, pET *NeRh50*"4, 6, pET *NeRh50*"3, 4, 6, and pET *NeRh50*"2, 3, 4, 6 has been successfully cloned in to

the pET expression vector. In the next part, the expression of the recombinant proteins were confirmed using specific antibodies (Table 2.1) (Table 3.2), western blotting, Flow cytometry, and TEM.

3.6.2 Confirming the presence of constructs by testing for proteins expression

The principle of identifying blood group specific antibodies in human sera by using their soluble antigens was successfully established for several blood groups such as; ABO, Lewis, P and I (Spitalnik *et al.*, 1983; Cowles & Blumberg, 1987; Daniels, 2002).

Recombinant proteins were successfully reported to express on cell membranes or as a soluble form and to exhibit blood-group-specific antigen activity. In 2007, there was a successful study by Ridgwell *et al.*, on the activity of the soluble recombinant proteins with Kell, Duffy and Lutheran blood group antigens. Moreover, Knops, Cromer, Chido/Rodgers, Indian and JMH systems were shown to be antigenically active by haemagglutination inhibition (Williams *et al.*, 1981; Telen *et al.*, 1994; Moulds & Rowe, 1996; Daniels, 2002)

In spite of the clinical significance of the Rh antigens, no studies have been published describing successful production of recombinant antigens for it, due to the highly conformation-dependent structure of the human Rh antigens which make it is difficult to construct the recombinant Rh antigen for the Rh protein *in vitro*.

The recombinant RhD may be helpful to further investigate the molecular basis and function of human RhD protein and could be applicable for screening anti-D in patient serum. It may also

form the technical platform to express all Rh system antigens including uncommon variants to be capable of detecting these antibodies in patient serum.

Based on the successful result of SDM work on expressing the recombinant RhCE in the K562 cell line, constructing the recombinant *RHCE* gene by replacing the human *RhCE* external loops (3, 4, and 6) with the corresponding, human *RhD* external loops (3, 4, and 6) sequences to construct a recombinant form of Rh protein, which has a backbone of RhCE and external loops from human RhD by Liu *et al.*, (1999 a, b) we have chosen the external loops (3, 4, and 6) of the human RhD protein to study. Furthermore, choosing the *E. coli* bacterial membrane to be the host of this experiment is because of the ease to grow it in the laboratory. Additionally, the topology similarities between the 12 membrane domains of the *N. europaea* Rh50 like protein (Figure 1.13) and the human Rh protein was the main reason for using this construct in this project to discover new information about Rh protein structures, and their antigens, unlike the *E. coli* AmtB protein (Figure 1.11), which has only 11 membrane domains. Thus, the *NeRh50* like protein could be used as a template to express human Rh epitopes on the extracellular surface of *E. coli*. Rather than express the hybrid constructs in *N. europaea* they were expressed in *E. coli* for ease of use (Cherif-Zahar *et al.*, 2007).

In this study *NeRh50* was used as a backbone combined with

human *RHD* external loops to create recombinant constructs in order to produce D-epitopes. This is a a fundamental difference compared to the work of Liu *et al.*, (1999). Hence by using *NeRh50* rather than RhCE as in Liu *et al.*, (1999), we can determine how the recombinant protein gives rise to the D epitopes. The assembly and expression of the *NeRh50* protein on the membrane does not require RHAG.

Furthermore, Liu *et al.*, (1999) work was based on expressing the recombinant construct *RHCE-RHD* in the K562 cell line, which is a human myelogenous leukaemia cell line, so, in that study they worked with cell lines naturally expressing RhAG and that is another fundamental change because this study is based on using *E. coli*, which have AmtB protein instead, to detect the expression of the recombinant gene *NeRH50"RHD* using specific Anti-D (Table 2.1). The advantage of using prokaryotic system (*E. coli*) over the eukaryotic system (K562 cell line) is, to express D-epitopes in system lacking of *RHAG* inorder to study the chance of being able to produce D-epitopes without the presence of this gene.

3.6.2.1 Western blot results

As part of this project, a custom antibody against human RhD external loop 6 (2A4/ 5G9) was produced from Dundee Cell Products. Moreover, confirming the presence of the recombinant gene *NeRh50"human RhD external loop 6* by detecting

the RhD external loop 6 protein expression on the *E. coli* membrane in; pET*NeRh50*"6, pET *NeRh50*"4, 6, pET *NeRh50*"3, 4, 6, and pET *NeRh50*"2, 3, 4, 6, constructs using the western-blot technique (Figure 3.31) (Figure 3.32).

Ghost membrane Rh⁺, Rh⁻ phenotypes, which were used as a control samples, they were not showing a direct reactivity with the peptides monoclonal anti-D loop 6 and to cover this issue, these samples were treated with the bromelain (Avent *et al.*, 1992) before run on the acrylamide gel to digest the external loop 6 and open it, which be easier to detect (see 2.2.2.4.7.2). Moreover, (2A4 and 5G9) anti-D which are able to detect the human RhD external loop 6, was expressed in all recombinant constructs as all constructs has the human RhD external loop 6 (Figure 3.31) (Figure 3.32).

Ghost membrane Rh⁺ phenotype was showed a faint single band around 32 kDa, while all recombinant constructs showed a single band around 37 kDa. Ghost membrane Rh⁻ phenotype and *NeRh50* wild type showed no band as expected due to the absence of the human RhD external loop 6.

Different band intensity between 2A4 and 5G9 shows that, 2A4 is a more effective clone of this anti-RhD loop 6 antibody than 5G9 and it is able to detect a higher expression of the human RhD external loop 6 in the ghost membrane Rh⁺ phenotype and in all constructs.

Anti-D loop 4 (2.2B9) was used to confirm the presence of the human RhD external fourth loop in recombinant constructs and control samples. It shows double bands with the positive samples the upper band is a background band and it does not related to RhD. The lower band, is the band of interest has a size approximately 32 kDa in ghost membrane Rh⁺ phenotype and the RhD30 wild type samples, while it has a size of 36 kDa in constructs samples (Figure 3.30). Furthermore, ghost membrane Rh⁻ phenotype showed a single band around 34 kDa, while *NeRh50* wild type and *NeRh50*"6 showed a single band close to 37 kDa, which represent the background bands due to the absence of human RhD external loop 4.

NeRh50 wild type is a glycoprotein, which contains an oligosaccharide chain attached to the external loop 1 and this give it a higher molecular weight than human RhD, which does not contain the carbohydrate chain. The difference in protein structure between ghost membrane sample with an Rh⁻ phenotype vs *NeRh50* wild type leads to the variation seen in band size when the 2.2B9 antibody was used for Western blotting. The difference between the 37 kDa band size for *NeRh50* wild type and 34 kDa band size for the Rh⁻ ghost membrane sample is due to the additional oligosaccharide chain in the former protein.

Anti-D loop 3, 4, 6 (MS-26) (Figure 3.34), shows no reaction with all constructs or controls in the western blot technique. The lack of an obvious band reflects that this MAb may recognise

a conformation- dependent antibody, not detectable by western blotting.

Anti-NeRh50 (1918, and 1917) (figure 3.35) (figure 3.36) both of these antibodies gave negative results with human ghost membrane Rh⁺, Rh⁻ phenotyped samples due to the absence of the *NeRh50* protein, while they show a clear band around 37 kDa for the *NeRh50* wild type sample as well as in *NeRh50*"6, *NeRh50*"4, 6, *NeRh50*"3, 4, 6, and *NeRh50*"2, 3, 4, 6. These results clearly demonstrate that the *E. coli* pET expression system works well and all constructs are successfully integrated into the *E. coli* inner membrane.

The variation between the positive band size between the human ghost membrane Rh⁺ phenotype sample (32 kDa) and all constructs (37 kDa) is due to the difference in protein structure backbone of the construct NeRh50 protein, a glycoprotein, which contains an oligosaccharide chain attached to the external loop 1 and this give it a higher molecular weight than human RhD, which does not contain the carbohydrate chain.

Furthermore, human RhD30 wild type and *NeRh50* wild type samples were used as controls to detect *NeRh50* external loop 1/ C-terminus and human RhD external loops.

According to Roudkenar *et al.*, (2008) they successfully detected the expression of cord RhD with Human anti-Rh (D) polyclonal antibody (Immunodiagnostika, USA) around 35 kDa, which confirm the band size was shown with anti-D antibodies against

human RhD external loops 4 and 6 in the ghost membrane Rh⁺ phenotype and in four constructs.

According to Cherif-Zahar *et al.*, (1991) they have His- tagged the C-terminus of *NeRh50* protein to able to detect *NeRh50* expression in the *E. coli* membrane with anti-His tag antibody by western blot at 37 kDa.

According to Apoil *et al.*, (1997) they were successfully blot the RhD protein with 33 kDa band size by using LOR15C9 Mab, which is an important antibody to distinguish between RhD and RhCE proteins.

3.6.2.2 Flow cytometry results

In 1999 a, Liu *et al.*, used monoclonal anti-D, anti-E, and anti-c antibodies, which were available through the Third International Workshop against human red blood cells and related antigens. Moreover, they successfully constructed a hybrid recombinant protein based on *RhCE* as a backbone and replaced the third, fourth, and sixth external loops with the corresponding external loops from the human *RhD*. Then, were able to screen the expression of all constructs with monoclonal anti-D, anti-E, and anti-c antibodies by the flow cytometry technique on the K562 cell line.

In this study, sample preparation and staining protocol took a long time and several repeat trials to optimize the correct modified method. Additionally, all construct samples were lysozyme

treated prior to staining them with the primary anti-D or anti-*NeRh50* antibodies (see 2.2.2.4.9.2) in order to make some holes in the *E. coli* membrane, and this also required optimisation of the centrifugation time (Soriano *et al.*, 2002) moreover, assess detecting the expression of the recombinant protein, which located on the inner *E. coli* membrane.

Table 3.2 shows monoclonal anti-D antibodies were used to detect the expression of human RhD external loops (2, 3, 4, and 6). Moreover, anti-*NeRh50* antibody were used to detect the expression of *NeRh50* external loop 1 and C-terminus.

In this study, all constructs show variation in FITC fluorescence values among different antibodies. The disadvantage of using *NeRh50* over human RhCE as a back bone in Liu *et al.*, 1999, is the difference in the expression pattern, it shows a higher expression pattern with using the human RhCE backbone.

*NeRh50"*3, 4, 6 construct (Table 4.3) showed the lowest expression with anti-D while it shows a moderate mean FITC fluorescence value with the anti-*NeRh50* comparing to control samples. Anti-D BRAD-3, which is specific for EpD 13 shows the highest value with RhD30 wild type control. Furthermore, in Liu *et al.*, 1999 BIRMA-D6 gave the highest mean value with recombinant RhCE protein with human RhD fourth and sixth external loop, while in this study BIRMA-D6 show the highest value with *NeRh50"*2, 3, 4, 6 construct.

3.6.2.3 TEM results

Thin layer of *E. coli* cells containing one of four constructs; pET*NeRh50*"6, pET *NeRh50*"4, 6, pET *NeRh50*"3, 4, 6, and pET *NeRh50*"2, 3, 4, 6, were used in immuno-gold staining technique. The human RhD external loop 4 and loop 6 expression were detected by anti-D (Table 3.2) 2.2B9 (Figure 3.42)/ 2A4 (Figure 3.43). There was a considerable change in the expression level between the antibodies used (Table 3.2); *NeRh50*"6 construct showed no binding and no gold trace with anti-D 2.2B9 (Figure 3.42), due to the absence of human RhD external loop 4 in this construct.

Moreover, all constructs showed high expression level with anti-D loop 6 2A4 (Figure 3.43), as was expected because loop 6 was present in all constructs, similar to the expression on the sample was stained with anti. *E. coli* (positive control) and higher than the expression of the samples were stained with anti-D 2.2B9. Furthermore, all constructs showed intermediate expression level anti-*NeRh50* 1918 (Figure 3.44) comparing to the positive control and to other anti-D antibodies.

Different dot size and intensity in images taken by TEM was due to the variation of the protein concentration in each sample: big and dark gold trace shows higher protein concentration.

To summarize this chapter, the recombinant *NeRh50''RHD* gene was successfully cloned and expressed. The expected size of recombinant *NeRh50''* human RhD external loops protein was detected and confirmed by western blot.

Flow cytometry and TEM imaging confirmed the existence of the human RhD external loops in the recombinant protein *NeRh50''* human RhD external loops. RhD antibody reacted with recombinant RhD antigen as well as with RhD30 wild type.

Chapter 4

Discussion and Conclusion

4.1 General discussion

Early in 1901, blood groups were discovered (Landsteiner, 1901; Landsteiner & Levine, 1927; Anstee *et al.*, 2010; Avent, 2016), and there have been considerable number of studies published in this field, reflecting the importance of this science in Immunohaematology. The difference in blood group phenotype remains a significant question and it is still attracting many researchers to study and discover more about different blood groups and their relationship to transfusion and /or transplantation rejection.

The principle of using the blood group antigens to identify blood group antibodies in human serum, is well established, although, this is limited by the rare occurrence of some blood group antigens. Blood group antigens can be replaced by recombinant forms replacing the requirement for the use of human cells, additionally it offers the ability to generate a rare blood group antigens. Furthermore, this has been achieved for many blood group systems: recombinant antigens were successfully produced for MNS blood group system by constructing a recombinant glycophorin A (GPA) carrying blood group M and N antigens, which were able to detect anti-GPA antibodies (Blackall *et al.*, 1992; Blackall *et al.*, 1994). In 2000, Yazdanbakhsh *et al.*, were able to express Knops, Kell and Duffy blood group antigens in a mouse erythroleukaemic cell line (MEL) and used them to

detect and identify human antibodies. In 2002, Jaskiewicz *et al.*, succeeded in expressing a recombinant Gerbich blood group antigen on the surface of wild type Chinese hamster ovary cells (CHO) and a monkey kidney cell line (COS 7) by using anti-GPC murine monoclonal antibodies.

Furthermore, Ridgwell *et al.*, (2007) looked at the activity of soluble recombinant proteins with Kell, Duffy and Lutheran blood group antigens and concluded that the soluble recombinant proteins can mimic the blood group antigens, and these proteins can react with the human-specific antibodies. Moreover, these kinds of proteins can be used to develop solid-phase, high-throughput blood group antibody screening and identification platforms. The *RH* blood group system has a vital importance as the second most important blood group system in transfusion medicine (Levine *et al.*, 1941; Avent, 2018). Anti-Rh system antibodies are major causes of transfusion reactions and haemolytic disease of the fetus and newborn (Landsteiner & Wiener, 1941; Race & Sanger, 1975; Mollison *et al.*, 1987; Van Kim *et al.*, 2006). The rapid detection of anti-Rh is of significant clinical relevance and is only performed using appropriate phenotyped red cells (Chang & Siegel, 1998). This has significant drawbacks in that certain Rh antigens are very rare and require the storage of a diminishing supply of red cells that must be stored in liquid nitrogen and are scarcely available (Chaudhari, 2007; Avent, 2018).

In spite of the clinical significant *RH* antigens and the considerable knowledge of their proteins (Levine *et al.*, 1941; Matassi *et al.*, 1999; Avent *et al.*, 2006; Van Kim *et al.*, 2006; Avent *et al.*, 2012; Mustafa *et al.*, 2014) efforts to produce recombinant Rh antigens have to date been unsuccessful in prokaryotes (Huang & Peng, 2005), due to the highly conformation-dependent structure of the human Rh protein which makes it difficult to construct.

Liu *et al.*, (1999 b), showed successful production of a human recombinant gene *RHCE-RHD* and a successful expression in the K562 cell line, a eukaryotic tissue culture system, taking into consideration the homology between *RHCE* and *RHD*, which are both are located on the same locus on chromosome 1p34-1p36 (Cherif-Zahar *et al.*, 1991; Matassi *et al.*, 1999). Both genes share comparable exon/intron organisation (Matassi *et al.*, 1999; Westhoff, 2004) (Figure 1.3). This similarity raised the probability of these recombinant constructs to successfully produce human D-epitopes, moreover, they detected the expression of human recombinant gene *RHCE-RHD* on eukaryotic system which is able to express RhAG, which is not the case with using a prokaryotic expression system (Huang & Peng, 2005) and as previously mentioned that the expression of RhD/RHCE antigens is depend on the presence of *RHAG* (Chérif-Zahar *et al.*, 1996; Avent *et al.*, 2000).

However, the knowledge about Rh proteins would transform the

ability to test for particular rare Rh antigens (of which there are 50 known) (Conroy *et al.*, 2005). Furthermore, a suitable expression system would yield significant new information regarding the topology of Rh epitopes (Mollison *et al.*, 1987; Eysers *et al.*, 1994). Recombinant Rh antigens could be used to rapidly and reliably detect anti-Rh in serum, and mono-specific antigens would greatly help resolve complex mixtures of antibodies in multi-transfused patients.

Therefore, the aim of the project was to construct a hybrid *N. europaea*-human *RHD* gene that is capable of being transcribed, translated and correctly inserted into the *E. coli* membrane, in order to develop novel diagnostic assays able to detect Rh antigens. *N. europaea* possesses an Rh50 protein which is similar in the transmembrane domain structure to human RhD and RhCE, with twelve transmembrane domains. Different regions of the human RhD and RhCE proteins were utilised to create these hybrids.

The topology similarities in membrane domains between the 12 transmembrane domains of the *N. europaea* Rh50 like protein (Figure 1.13) and the human Rh protein was the main reason of using this type of bacteria in this project to discover new information about Rh protein structures, and their antigens, unlike the *E. coli* AmtB protein (Figure 1.11), which only has 11 membrane domains. Thus the *NeRh50* like protein was used as a backbone for the human Rh epitopes as *NeRh50* can be

expressed in *E. coli* (Cherif-Zahar *et al.*, 2007).

Furthermore, the objective of this research was to use the *NeRh50* human RhD external loops protein as a template to express human RhD epitopes on the bacterial membrane. Coupled with considerable knowledge regarding the predicted molecular topology of Rh antigen epitopes as defined by site directed mutagenesis (SDM) (Liu *et al.*, 1999 b), recombinant *Ne-human Rh* proteins may provide a suitable alternative to produce vast amounts of recombinant Rh antigens in a bacterial system.

Based on the successful result of SDM work in constructing the recombinant *RHCE* gene by Liu *et al.*, 1999 b, we chose the external loops (6, 4, and 3) of the human RhD protein to study. Loop 6 was chosen to start with because of its location between two highly conserved transmembrane domains (11-12), which increased the probability of this sequence being expressed successfully by *N. europaea*. Following working with loop 6, loops 2, 3, and 4 were also chosen.

Using recombinant antigens to detect the anti-D in patient serum will be a useful tool in cases such as, pregnant women carrying an RhD positive fetus, in multi-transfused patient, in individuals with partial D and or weak D.

Moreover, it possible to follow the same method and construct a hybrids mimicking partial D in order to accurately phenotype individuals carrying variant antibodies and hence reduce mismatching. Although the importance of this approach, there are

some limitations such as the possibility of the carbohydrate parts of *E. coli* membrane could copy human A/ B antigens, which can show some false positive results. This limitation can be solved by using the L-form bacteria (Kujau *et al.*, 1998; Rippmann *et al.*, 1998; Errington, 2017). Moreover, other limitation is having a variation in the expression pattern of different batches by using flow-cytometry.

4.2 Expression of D-epitope in *E. coli*

E. coli was used as the expression system due to the simplicity of using *E. coli* in the laboratory and it is widely used as a model and host organism in biotechnology and microbiology, especially when working with recombinant DNA (Lee, 1996; Makrides 1996).

In this study, the idea was that the bacteria express just one Rh-D antigen, which has many epitopes, although bacterial carbohydrate structures may mimic red cell antigens. Thus, the need of having recombinant antigen was developed in this field to cover these problems.

The recombinant *NeRH50''RHD* gene with different human RhD external loops (2, 3, 4, 6) to create: *NeRh50''6*, *NeRh50''4*, *6*, *NeRh50''3*, *4*, *6*, and *NeRh50''2*, *3*, *4*, *6*, were successfully cloned in to the pET expression vector and D epitope was successfully

expressed in *E. coli*. Western blotting, flow cytometry analysis, and TEM techniques were used to determine the expression of human D-epitopes in the *E. coli* membrane by using anti-D and anti-*NeRh50* antibodies (Table 2.1).

The expected size of recombinant *NeRh50* human RhD external loops protein was 32-37 kDa. According to some studies on human RhD, they were able to detect it by western blotting at 30-35 kDa using monoclonal anti-D (Moore *et al.*, 1982; Gahmberg, 1982, 1983). That concludes, D-epitopes in the recombinant protein have been successfully detected and confirmed by western blotting by use the anti-D in this study (Figure 3.30 - 3.36). The expected size range of recombinant *NeRh50* human RhD protein was higher than 30-35 kDa, which is the range of human D-epitope. This was due to the backbone of the recombinant protein used in this study being *NeRh50* wild type is a glycoprotein, which contains an oligosaccharide chain attached to the external loop 1 and this gives it a higher molecular weight than human RhD, which does not contain the carbohydrate chain. Different FITC- Fluorescence flow cytometry mean values analysis, and TEM imaging confirmed the existence of the human RhD external loops in the recombinant protein (*NeRh50* human RhD external loops). RhD antibody reacted with recombinant *NeRh50* RhD antigen as well as with RhD30 wild type, confirming the presence of the D-epitopes in the recombinant protein. Furthermore, a specific antibody raised against the ex-

ternal loop1 and C-terminus peptides of *NeRh50* protein was ordered from Dundee Cell Product. This antibody was used to confirm the existence of the *NeRh50* protein in *E. coli*.

This project has significantly advanced from the Liu *et al.*, 1999a and b studies. In their study they used the human *RHCE* as a backbone, while in this study *NeRH50* was used to be the backbone of the recombinant construct. Moreover, this study showed successful detection of D-epitopes from the recombinant protein *NeRh50*"human RhD external loops, which proves that the recombinant protein expresses D epitopes even though only the second, third, fourth, and sixth human RhD external domains have been added to the *NeRh50* backbone.

RhD was successfully expressed in K562 cell line (Yan *et al.*, 2005) moreover, recombinants *RHCE-RHD* were expressed in K562 cell line (Liu *et al.*, 1999). The K562 cell line is a eukaryote expression system and dependent on the presence of RhAG. Although the prokaryotic expression system represented in *E. coli* does not require the presence of RhAG, it successfully expressed D-epitopes from the recombinant *NeRh50*"human RhD external loops.

Finally, so far, no study has been published describing a successful expression of a recombinant RhD in a bacterial system and this project seems to be the first one. Furthermore, the antigenicity of the recombinant protein *NeRh50*"human RhD external loops was confirmed by western blotting, flow cytometry

analysis and TEM by using anti-D and anti-*NeRh50* antibodies. Moreover, this study proved some points; firstly, D-epitope expression does not require the presence of RhAG and the external predicted domains on the RhD protein can alone be responsible for expression. Secondly, using the correct expression system, Rh antigens can be expressed. A false positive and/or a false negative reaction could be seen from using a recombinant antigens with human sera and for that many trials and modification should be required.

Moreover, using a recombinant of the bacterial membrane domains coupled with humanized extracellular regions will help to discover new features about RhD protein and to investigate the molecular basis of RhD protein also to discover the main function of this protein.

4.3 Conclusions

In this project different techniques, western blots, flow cytometry and transmission electron microscopy have been used to detect expression of recombinant constructs of *NeRh50* human RhD external loops in the *E. coli* inner membrane combined with the using of different antibodies (Table 3.2).

2.2B9 which is a mAb used to detect the expression of human RhD external loop 4, has been used in all techniques and it shows negative results with the recombinant *NeRh50* as ex-

pected, as this construct is missing the human 4th external loop. Moreover, this mAb showed positive results with the other recombinant constructs.

2A4 is a mAb used to detect the expression of human RhD external loop 6 showed positive results with all constructs in three techniques. Moreover, it shows the highest number of black dots comparing to other antibodies with all constructs. MS26 is a mAb used to detect the expression of human RhD external loop 3, 4, 6. Although it was not able to detect any expression in either constructs or control samples by western blots but it was showed different expression pattern with flow cytometry.

4.4 Future Work

Firstly, the sequence of the human RhD external loops needs to be confirmed by designing new primer sets and resequencing all constructs using Sanger sequencing or NGS.

Secondly, it would be valuable to generate a custom antibody against human RhD second external loop. There have been no studies working with this loop before and the antibody would help to detect the second loop and distinguish between the recombinant constructs; *NeRh50''6*, *NeRh50''4*, *6*, *NeRh50''3*, *4*, *6*, and *NeRh50''2*, *3*, *4*, *6*.

Thirdly, grow a stable L-form *E. coli* cell strains, which are a wall-less, or wall deficient cells (Kujau *et al.*, 1998; Rippmann *et al.*, 1998; Errington, 2017) to use as an expression system for the recombinant gene (*NeRh50''RhD*) and that is because Rh antigens are expressed on the *E.coli* inner membrane, and thus any recombinant antigen system would need to remove *E. coli* external membrane.

Knocking down the AmtB gene in the L-form *E. coli* and put the recombinant construct and try to grow them with the expression from the recombinant protein, thus will able to investigate more about the RhD function.

There are over 100 monoclonal anti-D and some of them are commercially available. Some of them have been used in this study (Table 2.1) to detect the human D-antigens and so, more

anti-D need to be used to screen all colonies to confirm the presence of D-epitopes in each construct. Furthermore, LOR-15C9 human mAbs anti-D (Zhu *et al.*, 1999), which is a loop 6 dependent Ab and can distinguish between RhD and RhCE, could be used to confirm if the RhD and RhCE proteins are expressed successfully on the *E. coli* membrane. Furthermore, wild type RhD30 clone would be used as a reference in the expression pattern because this clone should express all D-epitopes.

After the successful work in expressing recombinant Rh antigens in this project, which were confirmed with Western blotting, flow cytometry analysis, and TEM techniques; it extends the possibility of having more mutant Rh clones by using CE-epitopes and try to express Rh (C, c, E, e) antigens in prokaryotic system, using anti-E and anti-C antibodies. Moreover, testing these recombinants in human sera will be one of the most practical procedures to investigate the efficiency of these recombinants.

Finally, enzyme-linked immunoassays (ELISA) technique could be used to assess the antigenicity of the expressed recombinant protein. Moreover, make it possible to detect anti-D in patient serum by immobilized *E. coli* membrane, which contain Rh antigens on a solid surface, then add the patient serum to detect their Rh antibodies. The most important element of the detection strategy is a highly specific antibody-antigen interaction.

Recently, there are progression and improvement in the genetic tools that lead to more understanding of the *E. coli* organism,

which leads to raise the importance of using *E. coli* as expression system of complex eukaryotic proteins. Moreover, it increases the ability to express and purify the desired recombinant protein without the need for animal models (Rosano & Ceccarelli, 2014).

References list

Agre, P., Smith, B. L., & Hartel-schenk, S. (1990). Biochemistry of the Erythrocyte Rh Polypeptides: A Review. *The Yale Journal of Biology and Medicine*, 63, 461–467.

Alloisio, N., Texier, P., Forissier, A., Ribeio, M.L., Morle, L., Bozon, M., Bursaux, E., Maillet, P., Tanner, M.J.A., Tamagnini, G. & Delaunay, J. (1993). Band 3 Coimbra: a variant associated with dominant hereditary spherocytosis and band 3 deficiency. (Abstract). *Blood*, 82, (Suppl. 1). 4a.

Alberts, B., Johnson, A., Lewis, J., Walter, P., Raff, M., Roberts, K.(2002). "Chapter 24". *Molecular Biology of the Cell* (4th ed.). Routledge. ISBN 978-0-8153-3288-6.

Anstee, D. J., East, S.,& Founder, A. (2010). Review article The relationship between blood groups and disease, 115(23), 4635–4643. <http://doi.org/10.1182/blood-2010-01-261859>.

Anthea, M., Hopkins, J., McLaughlin, W. C., Johnson, S., Quon-Warner, M., LaHart, D., Wright, J. D. (1993). *Human Biology and Health*. Englewood Cliffs, New Jersey, USA: Prentice Hall. ISBN 0-13-981176-1.

Arp, D., Sayavedra-Soto, L., & Hommes, N. (2002). Molecular biology and biochemistry of ammonia oxidation by *Nitrosomonas europaea*. *Archives of Microbiology*, 178(4), 250–255. <http://doi.org/10.1007/s00203-002-0452-0>.

Arce, M.A., Thompson, E.S., Wagner, S., Coyne, K.E., Ferdman, B.A., Lublin, D.M. (1993). Molecular cloning of RhD

cDNA derived from a gene present in RhD-positive, but not RhD-negative individuals. *Blood* 1993; 82:651–655.

Avice, M.N., Rubio, M., Sergerie, M., Delespesse, G., Sarfati, M. (2001). Role of CD47 in the induction of human naïve T cell anergy. *J Immunol* 167: 2459–2468.

Avent, N., Judson, P. A., Parsons, S. F., Mallinson, G., Anstee, D. J., Tanner, M. J., Evans, P. R., Hodges, E., Maciver, A.G., Holmes, C. (1988). Monoclonal antibodies recognize different membrane proteins that are deficient in Rh_{null} erythrocytes. One group of antibodies reacts with a variety of cells and tissues whereas the other group is erythroid-specific, *Biochemical Journal*, 251, 499–505.

Avent, N. D., Ridgwell, K., Tanner, M. J. A., Anstee, D. J. (1990). cDNA cloning of 30kDa erythrocyte membrane protein associated with the Rh (Rhesus)- blood-group antigen expression, *Biochemical Journal*, 271, 821–5.

Avent, N.D., Butcher, S. K., Liu, W., Mawby, W. Y., Mallinson, G., Parsons, S. F., Anstee, D. J., Tanner, M. J. A. (1992). Localization of the C termini of the Rh (rhesus) polypeptides to the cytoplasmatic face of the human erythrocyte membrane, *Journal of Biological Chemistry*, 267, 15134–9.

Avent, N. D., Reid, M. E. (2000). The Rh blood group system: a review Review article The Rh blood group system: a review, 95(2), 375–387.

Avent, N. D., Madgett, T. E., Lee, Z. E., Head, D. J., Mad-

docks, D. G., & Skinner, L. H. (2006). Molecular biology of Rh proteins and relevance to molecular medicine. *Expert Reviews in Molecular Medicine*, 8(13), 1–20. <http://doi.org/10.1017/S1462399406010969>.

Avent N.D. (2007). New insight into the Rh system: structure and function. *Vox Sanguinis ISBT Science series 2*: 35-43.

Avent, N.D. (2018). *Transfusion and Transplantation Science (Fundamentals of Biomedical Science)* . Oxford, United Kingdom: Oxford University Press.

Avene, M.A., Thompson, E. S., Wagner, S., Coyne, K.E., Ferdman, B. A., Lublin, D. M.(1993). Molecular cloning of RhD cDNA derived from a gene present in RhD-positive, but not RhD-negative individuals. *Blood* 1993; 82:651–655.

Bakouh, N., Benjelloun, F., Hulin, P., Brouillard, F., Edelman, A., Cherif-Zahar, B. & Planelles, G. (2004). NH_3 is involved in the NH_4^+ transport induced by the functional expression of the human Rh C glycoprotein. *Journal of Biological Chemistry*, 279, 15975–15983.

Bentley, R., Meganathan, R. (1982). Biosynthesis of vitamin K (menaquinone) in bacteria. *Microbiological Reviews*. 46 (3): 241–80. PMC 281544. PMID 6127606 .

Bennett, V., Baines, A.J. (2001). Spectrin and ankyrin-based pathways: metazoan inventions for integrating cells into tissues. *Physiol. Rev.* 81 (3): 1353–92.

Bhutani, V. K., Zipursky, A., Blencowe, H., Khanna, R.,

Sgro, M., Ebbesen, F., Bell, J., Mori, R., Slusher, T.M., Fahmy, N., Paul, V.K., Du, L., Okolo, A.A., de Almeida, M.F., Olu-sanya, B.O., Kumar, P., Cousens, S., Lawn, J.E. (2013). Neonatal hyperbilirubinemia and Rhesus disease of the newborn: incidence and impairment estimates for 2010 at regional and global levels. *Pediatric Research*, 74 Supply, 86–100. <http://doi.org/10.1038/pr.2013.208>.

Blackall, D.P., Ugorski, M., Smith, M. E., Pahlson, P., Spitalnik, S.L. (1992). The binding of human alloantibodies to recombinant glycophorin A. *Transfusion* 1992;32:629–32.

Blackall, D.P., Ugorski, M., Pahlson., P, Shakin-Eshleman, S.H., Spitalnik, S.L.(1994). A molecular biological approach to study the fine specificity of antibodies directed to the MN human blood group antigens. *J Immunol* 1994;152:2241–7.

Blakey, D., Leech, A., Thomas, G.H., Coutts, G., Findlay, K., & Merrick, M. (2002). Purification of the *Escherichia coli* ammonium transporter AmtB reveals a trimeric stoichiometry, *Biochem. J.* 364 527–535.

Bowman, J. M. (1775). RhD Hemolytic Disease of the Newborn. *The New England Journal of Medicine* , 339 (24).

Bowman J. M. (1988). The prevention of Rh immunization. *Trans Med Rev*; 2: 129–150.

Bruce, L. J., Beckmann, R., Ribeiro, M. L., Peters, L. L., Chasis, J. A., Delaunay, J., Mohanadas, N., Anstee, D. J. & Tanner, M. J. (2003). A band 3-based macrocomplex of integral

and peripheral proteins in the RBC membrane. *Blood*, 101, 4180-8.

Cartron, J. P. (1994), Defining the blood group antigens. *Biochemistry and molecular genetics, Blood reviews*, 8, 199-212.

Cartron, J. P. (1999). RH blood group system and molecular basis of Rh-deficiency, *Bailliere's Best Practice and Research Clinical Haematology*, 12, 655-89.

Chain, P., Lamerdin, J., Larimer, F., Regala, W., Lao, V., Land, M., Hauser, L., Hooper, A., Klotz, M., Norton, J. (2003). *J Bacteriol* 185:2759–2773.

Chang, T. Y., & Siegel, D. L. (1998). Genetic and immunological properties of phage-displayed human anti-Rh(D) antibodies: implications for Rh(D) epitope topology. *Blood*, 91(8), 3066–3078. [http://doi.org/10.1016/S0887-7963\(98\)80011-4](http://doi.org/10.1016/S0887-7963(98)80011-4).

Cheng, A., van Hock, A.N., Yeager, M.(1997). Three dimensional organization of a human water channel. *Nature* 387:627-630.

Cherif -Zaher, B., Van Kim, C., Blanchard, D., Bailly, P., Hermand, P., Salmon, C., Cartron, J. P., Coliny, Y. (1990). Molecular cloning and protein structure of a human blood group Rh polypeptide, *Proceedings of the National Academy of Science of the United States of America*, 87, 6243-7.

Cherif -Zaher, B., Raynal, V., Le Van Kim, C., D'Ambrosio, a M., Bailly, P., Cartron, J. P., & Colin, Y. (1993). Structure and expression of the RH locus in the Rh-deficiency syndrome.

Blood, 82(2), 656–62.

Cherif-Zahar, B., Raynal, V., Gane, P., Mattei, M. G., Bailly, P., Gibbs, B., Colin, Y., Cartron, J. P. (1996). Candidate gene acting as a suppressor of the RH locus in most cases of Rh-deficiency, *Nature Genetics*, 12, 168-173.

Cherif -Zaher, B., Matassi, G., Raynal, V., Gane, P., Mempel, W., Perez, C., Cartron, J. P. (1998). Molecular defects of the RHCE gene in Rh-deficient individuals of the amorph type, *Blood*, 92, 639-46.

Cherif-Zahar, B., Durand, A., Schmidt, I., Hamdaoui, N., Matic, I., Merrick, M., & Matassi, G. (2007). Evolution and functional characterization of the RH50 gene from the ammonia-oxidizing bacterium *Nitrosomonas europaea*. *Journal of Bacteriology*, 189 (24), 9090–9100.

Chaudhuri, A., Nielsen, S., Elkjaer, M. L., Zbrzezna, V., Fang, F. & Pogo, A. O. (1997). Detection of Duffy antigen in the plasma membranes and caveolae of vascular endothelial and epithelial cells of non-erythroid organs. *Blood*, 89, 701-12.

Chown, B., Lewis, M., Kaita, H., Lowen, B. (1972). An unlinked modifier of blood groups: effects when heterozygous and when homozygous, *American Journal of Human Genetics*, 24, 623-37.

Coombs, R. R. & Roberts, F. (1959). The antiglobulin reaction. *Br Med Bull*, 15, 113-8.

Colin, Y., Chérif-Zahar, B., Le Van Kim, C., Raynal, V.,

Van Huffel, V., & Cartron, J. P. (1991). Genetic basis of the RhD-positive and RhD-negative blood group polymorphism as determined by Southern analysis. *Blood*, 78, 2747–2752.

Conroy, M. J., Bullough, P., Merrick, M., & Avent, N. D. (2005). Modelling the human rhesus proteins: Implications for structure and function. *British Journal of Haematology*, 131(4), 543–551. doi:10.1111/j.1365-2141.2005.05786.

Cowles, J.W. & Blumberg, N. (1987). Neutralization of P blood group antibodies by synthetic solid-phase antigens. *Transfusion*, 27, 272–275.

Dahr, W., Kordowicz, M., Moulds, J., Gielen, W., Lebeck, L., Kruger, J. (1987), Characterization of the sialoglycoprotein and its antigens in Rh_{null} erythrocytes, *Blut*, 54, 13-24.

Davies, D.R., Cohen, G.H.(1996). Interactions of protein antigens with antibodies. *Proc Natl Acad Sci U S A* 93:7-12.

Daniels, G., Flegel, W. A., Fletcher, A., Garratty, G., Levene, C., Lomas-Francis, C., Moulds, J. M., Olsson, M. L., Overbeeke, M. A., Poole, J., Reid, M. E., Rouger, P., Van Der Schoot, C. E., Scott, M., Sistonen, P., Smart, E., Storry, J. R., Tani, Y., YU, L. C., Wendel, S., Westhoff, C. M. & Zelinski, T. (2007). International Society of Blood Transfusion Committee on Terminology for Red Cell Surface Antigens: Cape Town report. *Vox Sang*, 92, 250-3.

Daniels, G. (2002). *Human Blood Groups* (2nd ed). Blackwell, Oxford.

Delaunay, J. (1995). Genetic disorders of the red cell membranes. *FEBS Letters*. 369 (1): 34-37. doi:10.1016/0014-5793(95)00460-Q

Eder, A. F., and Chambers, L. A. (2007). Noninfectious Complications of Blood Transfusion. *Archives of Pathology & Laboratory Medicine*: May 2007, Vol. 131, No. 5, pp. 708-718.

Eyers, S. C., Ridgwell, K., Mawby, W. J., & Tanner, M. J. (1994). Topology and organization of human Rh (rhesus) blood group-related polypeptides. *Journal of Biological Chemistry*, 269(9), 6417–6423.

Errington, J. (2017): Cell wall-deficient, L-form bacteria in the 21st century: a personal perspective, *Biochem Soc Trans* . 45(2): 287–295.doi: 10.1042/BST20160435.

Faas, B. H., Simsek, S., Bleeker, P. M. M., Overbeeke, M. A. M., Cuijpers, H. TH. M., Von Dem Borne, A. E. G. KR.,Van Der Schoot, C. E. (1995). Rh E/e-genotyping by Allele Specific Primer Amplification (ASPA), *Blood*, 85, 829-32.

Fairbanks, G., Steck, T. L., & Wallachl, D. F. H. (1968). Electrophoretic Analysis. *Biochemistry*, 10(13), 2606–2617. doi:10.1021/bi00789a030.

Fisher, R. A., Race, R. R., Taylor, G. L. (1944). Mutation and the Rhesus reaction. *Nature*.153: 106. doi:10.1038/153106b0.

Flegel, W. A., Wagner, F. F. (1996). The frequency of RHD protein variants in Caucasians, *Transfusion Clinique et Biologique*, 3, 10S.

Gahmberg, C. G. (1983). Molecular characterization of the human red cell Rho(D) antigen. *EMBO J* 2:223-227.

Matassi, G., Chérif-Zahar, B., Pesole, G., Raynal, V., Cartron, J.P.(1999). The Members of the RH Gene Family (RH50 and RH30) Followed Different Evolutionary Pathways. *journal of molecular eveloution*, 48, 151-159.

Goldstein, J. (1989). Conversion of ABO blood groups. *Trans-fuse Med Rev* 1989;3:206–12.

Genetics Home Reference, (2014) 'Chromosomes'. [Online]. Available at: <http://ghr.nlm.nih.gov/chromosome/> (Accessed: December 3rd 2015).

Giangrande. P.L.F, (2000). The history of blood transfusion. <https://doi.org/10.1046/j.1365-2141.2000.02139>.

Gruswitz, F., Chaudhary, S., Ho, J.D., Schlessinger, A., Pezeshki, B., Ho, C.M., Sali, A., Westhoff, C.M., Stroud, R.M. (2010) . Function of human Rh based on structure of RhCG at 2.1 Å. *Proc.Natl.Acad.Sci. USA* 107: 9638-9643.

Hosoi, E., Rath, G., Mitra, R., & Mishra, N. (2014). Blood groups systems. *Indian Journal of Anaesthesia*. <http://doi.org/10.4103/0019-5049.144645>.

Huang, C.H. (1998). The Human Rh50 Glycoprotein Gene, 273(4), 2207–2213.

Huang, C. H., Chen, Y., Reid, M. E., Seidl, C. (1998). Rh_{null} Disease: The amorph type results from a novel double mutation in RhCe gene on D-negative background, *Blood*, 92, 664-671.

Huang, C.H., Liu, P.Z., Cheng, J.G. (2000). Molecular biology and genetics of the Rh blood group system. *Semin Hematol*; 37: 150–65.

Huang, C.H., Liu, P.Z. (2001). New insights into the Rh superfamily of genes and proteins in erythroid cells and non-erythroid tissues. *Blood Cells Mol Dis*. 27(1):90-101.

Huang, C. H., Peng, J.,(2005) : Evolutionary conservation and diversification of Rh family genes and proteins. *Proc Natl Acad Sci USA*; 102:15512–15517.

Iwamoto, S., LI, J., Omi, T., Ikemoto, S. & Kajii, E. (1996 a). Identification of a novel exon and spliced form of Duffy mRNA that is the predominant transcript in both erythroid and post capillary venule endothelium. *Blood*, 87, 378-385.

Iwamoto, S., LI, J., Sugimoto, N., Okuda, H. & Kajii, E. (1996 b). Characterization of the Duffy Gene Promotor: Evidence for Tissue-Specific Abolishment of Expression in Fy(ab) of Black Individuals. *Biochemical and Biophysical Research Communications*, 222, 852-859.

Iwamoto, S. (2005). Molecular aspects of Rh antigens. *Legal Medicine*, 7(4), 270–273. doi:10.1016/j.legalmed.2004.12.002.

Javelle, A., Lupo, D., Zheng, L., Dan Li, X., Winkler, F., and Merrick, M., (2006). An Unusual Twin-His Arrangement in the Pore of Ammonia Channels Is Essential for Substrate Conductance. *Journal of Biological Chemistry* 281 (51): 39492–98. doi:10.1074/jbc.M608325200.

Jones, J., Scott, M.L.& Voak, D. (1995). Monoclonal anti-D specificity and Rh D structure: criteria for selection of monoclonal anti-D reagents for routine typing of patient and donors. *Transfusion Medicine*, 5, 171-184.

Kaita, H., Lewis, M., Chown, B., Gard, E. (1959). A further example of the Kell blood group phenotype K-,k-,Kp(a-b-). *Nature*. 183 (4675): 1586.doi:10.1038/1831586b0.

Khademi. S., Connell III. J.O., Remis. J., Robles-Colmenares. Y., Miercke .L.J., & Stroud. R.M. (2004). Mechanism of ammonia transport by Amt/MEP/Rh: structure of AmtB at 1.35 Å, *Science* 305 1587–1594.

Kleiner, D. (1985). *FEMS Microbiol. Rev.* 32, 87–100.

Kustu, S., & Inwood, W. (2006). Biological gas channels for NH₃ and CO₂: evidence that Rh (Rhesus) proteins are CO₂ channels. *Transfusion Clinique et Biologique*, 13(1-2 SPEC. ISS.), 103–110. <http://doi.org/10.1016/j.tracli.2006.03.001>.

Kuypers, F., Van Linde-Sibenius -Trip, M., Roelosen, B., Tanner, M. J., Anstee, D. J., Den Kamp, J. A. (1984). Rh_{null} human erythrocytes have an abnormal membrane phospholipid organization, *Biochemical Journal*, 221, 931-4.

Kruskall, M.S., AuBuchon, J.P., Anthony, K.Y., Hersche, L., Pickard, C., Biehl R. (2000). Transfusion to blood group A and O patients of group B RBCs that have been enzymatically converted to group O. *Transfusion* 2000;40:1290–8.

Lauf, P. K., Joiner, C. H. (1976). Increased potassium trans-

port and ouabain binding in human Rh_{null} red blood cells, Blood, 48, 457-68.

Landsteiner, K. (1901). Ueber Agglutinationsersche inungen normalen menschlichen Blutes. About Agglutinationsersche inungen normal human blood. Wien. Klin. Wschr. 14, 1132. 6.

Landsteiner, K., Levine, P. (1927). A new agglutinable factor differentiating individual human bloods. Proc Soc Exp Biol, (N.Y.) 24, 600. 7.

Landsteiner, K., Levine, P. (1927). Further observations on individual differences of human blood. Proc Soc Exp Biol, 24, 941.

Landsteiner, K., Wiener, A. S. (1940). An agglutinable factor in human blood recognized by immune sera for rhesus blood. Proc Soc Exp Biol Med.;43:223.

Landsteiner, K., Wiener, A. S. (1941). Studies on an agglutino-gen (Rh) in human blood reacting with anti-rhesus sera and with human isoantibodies. J Exp Med 74:309-320.

Landsteiner, K.(1961). On Agglutination of Normal Human Blood. The journal of AABB.

Lang, F., Lang, E. & Foller, M. (2012). Physiology and pathophysiology of eryptosis. Transfus Med Hemother, 39, 308-14.

Levine, P. Stetson, R.E. (1939). An unusual case of intra-group agglutination. J Amer Med Ass; 113: 126–127.

Levine, P., Burnham, L., Katzin, E.M., & Vogel, P. (1941). The role of iso-immunization in the pathogenesis of erythroblastosis fetal is. *American Journal of Obstetrics and Gynecology*, 42: 925-937.

Levine, P. (1961). The Question of D (Rho) Antigenic Sites on Human Spermatozoa <https://doi.org/10.1111/j.1423-0410.1961.tb04060.x>

Lee, S.Y. (1996). High cell-density culture of *Escherichia coli*. *Trends in Biotechnology*. 14 (3): 98-105. Doi: 10.1016/0167-7799(96)80930-9 PMID8867291.

Liu, W., Avent, N. D., Jones, J. W., Scott, M. L., & Voak, D. (1999 a). Molecular configuration of Rh D epitopes as defined by site-directed mutagenesis and expression of mutant Rh constructs in K562 erythroleukemia cells. *Blood*, 94(12), 3986–96. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10590042>

Liu, W., Smythe, J. S., Scott, M. L., Jones, J. W., Voak, D., & Avent, N. D. (1999 b). Site-directed mutagenesis of the human D antigen: definition of D epitopes on the sixth external domain of the D protein expressed on K562 cells. *Transfusion*, 39(1), 17–25.

Lomas, C., Tippet, P., Thompson, K.M., Melamed, M.D. & Hughes-Jones, N.C. (1989). Demonstration of seven epitopes on the Rh D antigen D using human monoclonal anti-D antibodies and red cells from D-categories. *Vox Sanguinis*, 57, 261-264.

Lomas, C., McColl, K., & Tippet, P. (1993). Further com-

plexities of the Rh antigen D disclosed by testing category D-II cells with monoclonal anti-D. *Transfusion Medicine*, 3, 67-69.

Lögberg, L., Reid, M. E., & Zelinski, T. (2011). Human Blood Group Genes 2010: Chromosomal Locations and Cloning Strategies Revisited. *Transfusion Medicine Reviews*, 25(1), 36–46. doi:10.1016/j.tmr.2010.08.005.

Lupo, D., Li, X.-D., Durand, A., Tomizaki, T., Cherif-Zahar, B., Matassi, G., Merrick, M., Winkler, F. K. (2007). The 1.3-Å resolution structure of *Nitrosomonas europaea* Rh50 and mechanistic implications for NH₃ transport by Rhesus family proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 19303–19308. <http://doi.org/10.1073/pnas.0706563104>.

Ludewig, U. (2004). Electroneutral ammonium transport by basolateral rhesus B glycoprotein. *Journal of Physiology*, 559, 751–759.

Ludewig, U. (2006). Ion Transport versus Gas Conduction: Function of AMT/Rh-Type Proteins. *Transfus Clin Biol* 13(1-2):111–16.

Lindberg, F. P., Lublin, D. N., Telen, M. J., Veile, R. A., Miller, Y. E., Donis-Keller, H., Brown, E. J. (1994). Rh-related antigen CD47 is the signal transducer integrin-associated protein, *Journal of Biological Chemistry*, 269, 1567-70.

Le van, K. C., Mouro, I., Cherif-Zahar, B., Raynal, V., Cherrier, C., Cartron, J. P., Colin, Y. (1992). Molecular cloning and

primary structure of the human blood group RhD polypeptide. Proc Natl Acad Sci USA; 89:10925–10929.

Magnificus, R. (2004). A Study on Weak D and the Function of the Rh Complex in Red Blood Cells.

Marfatia, S. M., Lue, R. A., Branton, D., Chishti, A. H. (1994). In vitro binding studies suggest a membrane-associated complex between erythroid p55, protein 4.1, and glycophorin C. J Biol Chem. 1994;269:8631–8634.

Marini, A.M., Urrestarazu, A., Beauwens, R. & Andre, B. (1997). The Rh (Rhesus) blood group polypeptides are related to NH₄⁺ transporters. Trends in Biochemical Sciences, 22, 460–461.

Marini, A. M., Matassi, G., Raynal, V., Andre, B., Cartron, J. P., and Cherif-Zahar, B. (2000). The human Rhesus-associated RhAG protein and a kidney homologue promote ammonium transport in yeast. Nat. Genet. 26, 341–344.

Matassi, G., Chérif-Zahar, B., Pesole, G., Raynal, V., & Cartron, J. P. (1999). The members of the RH gene family (RH50 and RH30) followed different evolutionary pathways. Journal of Molecular Evolution, 48(2), 151–159.
<http://doi.org/10.1007/PL00006453>.

McTavish, H., Fuchs, J. a., & Hooper, A. B. (1993). Sequence of the Gene Coding for Ammonia Monooxygenase in *Nitrosomonas europaea*. Journal of Bacteriology, 175(8), 2436–2444.

Mustafa, M. H. I., Elmisbah, T. E., Salim, A.-M. M., &

Ahmed, M. A. M. (2014). RHD and RHCE frequencies and gene complexes among Major Tribes of Turabah Province , Saudia Arabia, (June), 573–578.

Mourant, E. A., Kopec, A. C., Domaniewska-Sobczak, K. (1976). The distribution of the human blood groups and other polymorphisms, 2nd edition, Oxford, Oxford University Press.

Mouro, I., Colin, Y., Cherif-Zahar, B., Cartron, J.P. & Le Van Kim, C. (1993). Molecular genetic basis of the human rhesus blood group system. *Nature Genetics*, 5, 62–65.

Mollison, P. L., Engelfriet, C. P. & Contreras, M. (1987). *Blood Transfusion in Clinical Medicine*, 8th ed., Blackwell, Oxford.

Moore, S., Green, C. (1987). The identification of specific Rhesus polypeptide blood group ABH active glycoprotein complexes in the human red cell membrane. *Biochem J* 244:735-742.

Moulds, J.M., & Rowe, K.E. (1996) Neutralization of Knops system antibodies using soluble complement receptor 1. *Transfusion*, 36, 517–520.

Miller, Y. E., Daniels, G. L., Jones, C., Palmer, D. K. (1987). Identification of a cell-surface antigen produced by a gene on chromosome 3 (cen-q22) and not expressed by Rh_{null} cells, *American Journal of Human Genetics*, 41, 1061-70.

Okuda H, Kawano M, Iwamoto S. (1997). The RHD gene is highly detectable in RhD-negative Japanese donors. *Journal of Clinical Investigation* 1997; 100: 373379.

Oldenborg, P. A., Zhelenyak, A., Fang, Y. F., Lagenaur, c. F., Gresham, H. D., Lindberg, F. P. (2000). Role of CD47 as a marker of self on red blood cells, *Science*, 288, 2051-4.

Pasini EM1,

Pasini, E. M., Kirkegaard, M., Mortensen, P., Lutz, H.U., Thomas, A.W., & Mann, M. (2006). In-Depth Analysis of the Membrane and Cytosolic Proteome of Red Blood Cells. *Blood* 108 (3): 791–801. doi:10.1182/blood-2005-11-007799.

Pollack, W., Gorman, J.G., Freda, V.J. (1968). Results of clinical trials of RhoGAM in women. *Transfusion* 8: 151.

Pinder,J.C .,Phethean,J.& Gratzer,W. B. (1978). *FEBS Lett.*92, 278-282.

Pilgrim, H., Lloyd-Jones, M., and Rees, A. (2009). Routine antenatal anti-D prophylaxis for RhD-negative women: a systematic review and economic evaluation, *Health Technology Assessment*; 13: 10.

Picchiassi, E., Di, C., Tarquini, F., & Centra, M. (2015). Non-Invasive Prenatal RHD Genotyping Using Cell-Free Fetal DNA from Maternal Plasma: An Italian Experience, 22-28. doi:10.1159/000370233.

Peng, J., Huang, C.H. (2006). *Transfus Clin Biol* 13:85–94.

Race, R. R., and Sanger, R. (1975). *Blood Groups in Man*, 6th Ed., Blackwell, Oxford.

Rath, G., Mitra, R., & Mishra, N. (2014). Blood groups systems. *Indian Journal of Anaesthesia*. doi:10.4103/0019-5049.14

Raetz, C. R. H. (1996). Bacterial lipopolysaccharides: a remarkable family of bioactive macroamphiphiles, p. 1035–1063. In F. C. Neidhardt et al, *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd. ed. American Society for Microbiology, Washington, D.C.

Reid.M.E, Mohandas.N. (2004). Red blood cell group antigens: Structure and function . *Semin Hematol* 41:93-117.

Reid, M. E., Lomas- Francis, C. & Olsson, M. L. (2012). 1 - Introduction. *The Blood Group Antigen Facts Book* (Third Edition). Boston: Academic Press. Rosen field RE, Allen FH Jr, Rubinstein P. (1973). Genetic model for the Rh blood-group system. *Proc Natl Acad Sci U S A*.70:1303-1307.

Rosenfield, R.E., Allen, F.H., Rubinstein, P. (1973). Genetic model for the Rh blood-group system. *Proc Natl Acad Sci U S A*.70:1303-1307.

Rosenfield, R.E., Allen, F.H., Swisher, S. N., Kochwa, S. (1979). Rh nomenclature. *Transfusion*.19:487.

Ripoche, P., Bertrand, O., Gane, P., Birkenmeier, C., Colin, Y. & Cartron, J.-P. (2004). Human Rhesus-associated glycoprotein mediates facilitated transport of NH₃ into red blood cells. *Proceedings of the National Academy of Sciences of the United States of America*, 101, 17222–17227.

Rippmann, J. F., Klein, M., Hoischen, C., Brocks, B., Rettig, W. J., Gumpert, J., Pfizenmaier, K., Mattes, R., Moosmayer, D.

(1998). Procaryotic expression of single-chain variable-fragment (scFv) antibodies: secretion in L-form cells of *Proteus mirabilis* leads to active product and overcomes the limitations of periplasmic expression in *Escherichia coli*. *Appl Environ Microbiol* 64(12):4862- 4869.

Ridgwell, K., Spurr, N. K., Laguda, B., MacGeoch, C., Avent, N. D., Tanner, M. J. (1992). Isolation of cDNA clones for a 50 kDa glycoprotein of the human erythrocyte membrane associated with Rh (rhesus) blood-group antigen expression. *Biochem J*; 287:223–228.

Ridgwell, K., Eyers, S. A., Mawby, w. J., Anstee, D. J., Tanner, M. J. (1994). Studies on the glycoprotein associated with Rh (Rhesus) blood group antigen expression in the human red blood cell membrane, *Journal of Biological Chemistry*, 269, 6410-16.

Ridgwell, K., Dixey, J., & Scott, M, L. (2007). Production of soluble recombinant proteins with Kell, Duffy and Lutheran blood group antigen activity, and their use in screening human sera for Kell, Duffy and Lutheran antibodies, *Transfusion Medicine*, 17, 384–394.

Schmidt, P. J., Vos, G. H. (1967). Multiple phenotypic abnormality associated with Rh_{null} (—/—), *Vox Sanguinis*, 13, 18-20.

Schmidt, P. J., Lostumbo, M. M., English, C. T., Hunter, B.(1969). Aberrant U blood group accompanying Rhnull *Trans-*

fusion,9, 40-2.

Scott, M.L., Voak, D., Jones, J.W., Avent, N.D., Liu, W., Hughes-Jones, N., Sonneborn, H. (1996). A structural model for 30 Rh D epitopes based on serological and DNA sequence data from partial D phenotypes. *Transfusion Clinique et Biologique*, 3, 391-396.

Singleton, P. (1999). *Bacteria in Biology, Biotechnology and Medicine*. (5th ed.).Wiley. pp. 444–454. ISBN 978-0-471-98880-9.

Singleton, B. K., Green, C. a, Avent, N. D., Martin, P. G., Smart, E., Daka, A., Daniels, G. (2000). The presence of an RHD pseudogene containing a 37 base pair duplication and a nonsense mutation in africans with the Rh D-negative blood group phenotype. *Blood*, 95(1), 12–18. [http://doi.org/10.1016/S0887-7963\(01\)80056-0](http://doi.org/10.1016/S0887-7963(01)80056-0).

Schwarz, H.P, Dorner, F. (2003). Karl landsteiner and his major contributions to haematology. 17–19, A-1220 Vienna, Austria.

Storry, J. R. & Olsson, M. L. (2009). The ABO blood group system revisited: a review and update. *Immunohematology*, 25, 48-59.

Sick, E., Jeanne, A., Schneider, C., Dedieu, S., Takeda, K., Martiny, L. (2012). CD47 update: a multifaceted actor in the tumour microenvironment of potential therapeutic interest. *Br. J. Pharmacol.* 167 (7): 1415–30.

Sloand, E., Kumar, P., Klein, H. G., Merritt, S., Sacher, R. (1994). Transfusion of blood components to persons infected with human immunodeficiency virus type 1: relationship to opportunistic infection. *Transfusion*.;34:48–53.

Spitalnik, S.L., Cowles, J.W., Cox, M.T. & Blumberg, N. (1983). Neutralization of Lewis blood group antibodies by synthetic immunoadsorbents. *American Journal of Clinical Pathology*, 80, 63–65.

Telen, M.J., Rao, N., Udani, M., Thompson, E.S., Kaufman, R.M. & Lublin, D.M. (1994). Molecular mapping of the Cromer blood group Cra and Tca epitopes of decay accelerating factor: toward the use of recombinant antigens in immunohematology. *Blood*, 84, 3205–3211.

Tills, D., Kopec, A. C., Tills, R. E. (1983). The distribution of the human blood groups and other polymorphisms, Oxford, Oxford University Press.

Tippett, P. & Sanger, R. (1962). Observations on subdivisions of Rh antigen, *D. Vox Sanguinis*, 7,9-13.

Tippett, P. & Sanger, R. (1977). Further observations on subdivisions of Rh antigen, *D. Arztl. Lab.* 23, 476-480.

Tournamille, C., Colin, Y., Cartron, J.P., Le Van-Kim, C. (1995). Disruption of a GATA motif in the Duffy gene promoter abolishes erythroid gene expression in Duffy-negative individuals. *Nat. Genet.* 10 (2): 224–8.

Thornton, J., Blakey, D., Scanlon, E., & Merrick, M. (2006).

FEMS Microbiol Lett 258:114–120.

Thomas, G.H., Mullins, J.G., Merrick, M. (2000). Mol Microbiol 37:331–344. 5.

Tortora, G. (2010). Microbiology: An Introduction. San Francisco, CA: Benjamin Cummings. pp. 85–87, 161, 165. ISBN 978-0-321-55007-1.

Tenaillon, O., Skurnik, D., Picard, B., & Denamur, E. (2010). The population genetics of commensal *Escherichia coli*. Nature Reviews. Microbiology. 8 (3):20717. doi:10.1038/nrmicro2298. PMID 20157339.

Urbaniak, S. J., & Greiss, M. a. (2000). RhD haemolytic disease of the fetus and the newborn. Blood Reviews, 14, 44–61. <http://doi.org/10.1054/blre.1999.0123>.

Van Kim, C. L., Colin, Y., & Cartron, J. P. (2006). Rh proteins: Key structural and functional components of the red cell membrane. Blood Reviews, 20(2), 93–110. <http://doi.org/10.1016/j.blre.2005.04.002>.

Von Wire, N., Merrick, M. (2004). Trends Curr Genet 9:95–120.

Von Heijne, G. (2006). Membrane-protein topology. Nature Reviews Molecular Cell Biology. 7 (12): 909–918.

Von Dem Borne, A. E., Bos, M. J., Lomas, C., Tippet, P., Bloy, C., Hermand, P., Cartron, J. P., Admiraal, L. G., Van De graaf, J., Overbeeke, M.A. (1990). Murine monoclonal antibodies against a unique determinant of erythrocytes, related to Rh and U antigens: expression on normal and malignant

erythrocyte precursors and Rh_{null} red cells, British Journal of Haematology, 75, 254-61.

Von, W., Merrick, M. (2004). Trends Curr Genet 9:95–120.

Von Heijne, G. (2006). Membrane-protein topology. Nature Reviews Molecular Cell Biology. 7 (12): 909–918.

Walz, T., Hirai, T., Murata, K. (1997). The three-dimensional structure of aquaporin 1. Nature 387:624-627.

Watson, J. D. (2007). Recombinant DNA: Genes and Genomes: A Short Course. San Francisco: W.H. Freeman. ISBN 0-7167-2866-4.

Westhoff, C. M. (2004). The Rh blood group system in review: A new face for the next decade. Transfusion, 44(11), 1663–1673. <http://doi.org/10.1111/j.0041-1132.2004.04237>.

Westhoff, C. M. (2005). Deciphering the Function of the Rh Family of Proteins. Transfusion 45(0041- 1132) LA - 121S.

Westhoff, C.M. & Wylie, D.E. (2006). Transport characteristics of mammalian Rh and Rh glycoproteins expressed in heterologous systems. Transfus Clin Biol. , 13 (1-2), 132.

Wiener, A. S. (1949). Genetics and Nomenclature of the Rh–Hr Blood Types. Antonie van Leeuwenhoek. Netherlands: Springer link. 15 (1): 17–28.doi:10.1007/BF02062626. ISSN 0003-6072. Retrieved 6 November 2010.

Wiener, D. I., Verlander, J.W. (2010). Molecular physiology of the Rh ammonia transport proteins.Curr Opin Nephrol Hypertens , 19 (5), 471–477.

Weidinger, K., Neuhauser, B., Gilch, S., Ludewig, U., Meyer, O., Schmidt, I. (2007). *FEMS Microbiol Lett* 273:260–267.

Williams, D., Johnson, C.L. & Marsh, W.L. (1981). Duffy antigen changes on red blood cells stored at low temperature. *Transfusion*, 21, 357–359.

Yawata, Y. (2003). *Cell Membrane: The Red Blood Cell as a Model*, Weinheim, WILEY-VCH.

Yamamoto, F. (2004). Review: ABO blood group system—ABH oligosaccharide antigens, anti-A and anti-B, A and B glycosyltransferases, and ABO genes. *Immunohematology / American Red Cross*, 20, 3-22.

Yazdanbakhsh, K., Oyen, R., Yu Q, Lee S, Antoniou, M., Chaudhuri, A., Reid, M. E. (2000). High-level stable expression of blood group antigens in a heterologous system. *Am J Hematol* ;63:114–24.

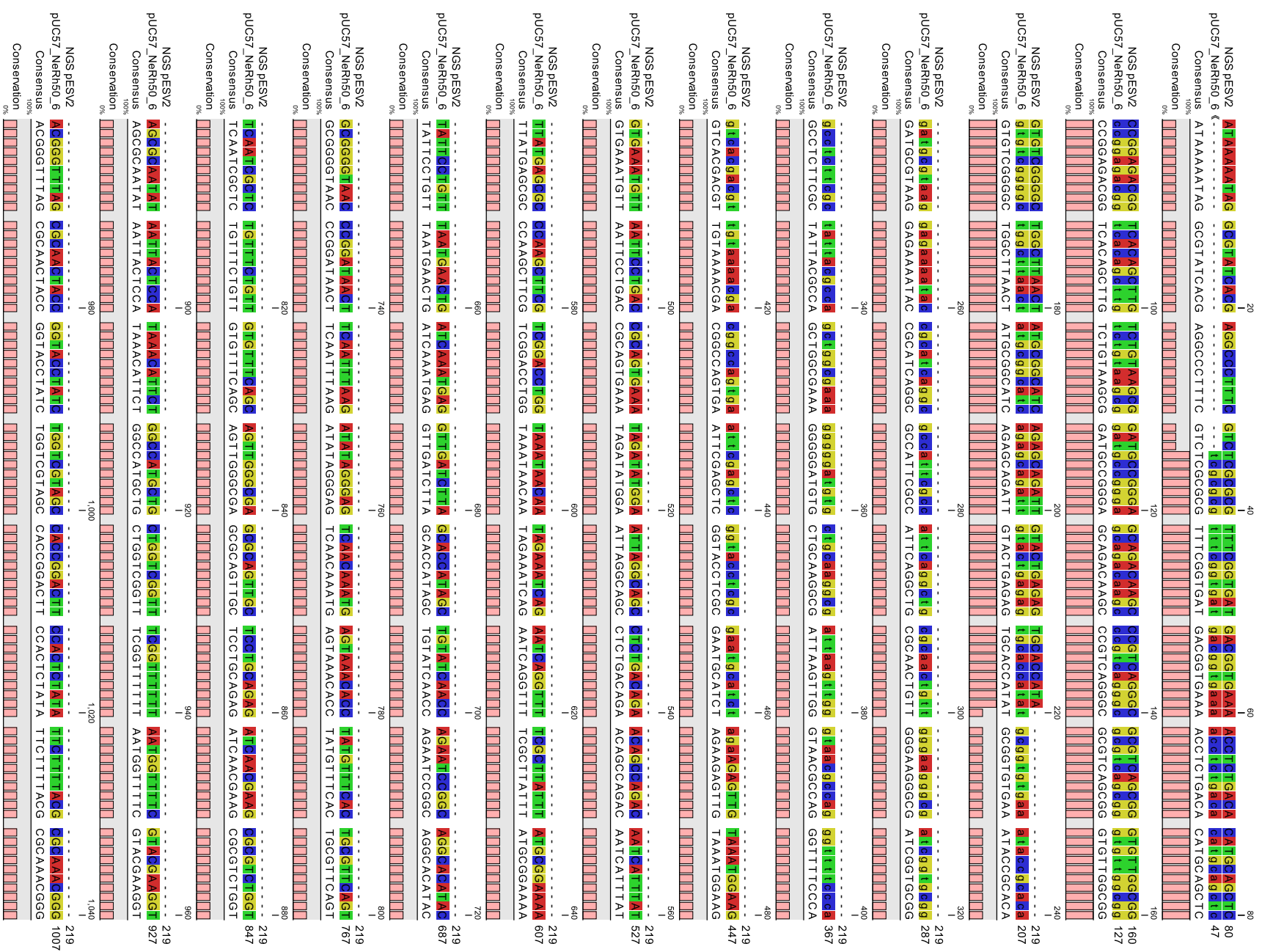
Zheng, L., Kostrewa, D., Berneche, S., Winkler, F.K. & Li, X.-D. (2004). The mechanism of ammonia transport based on the crystal structure of AmtB of *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*, 101, 17090–17095.

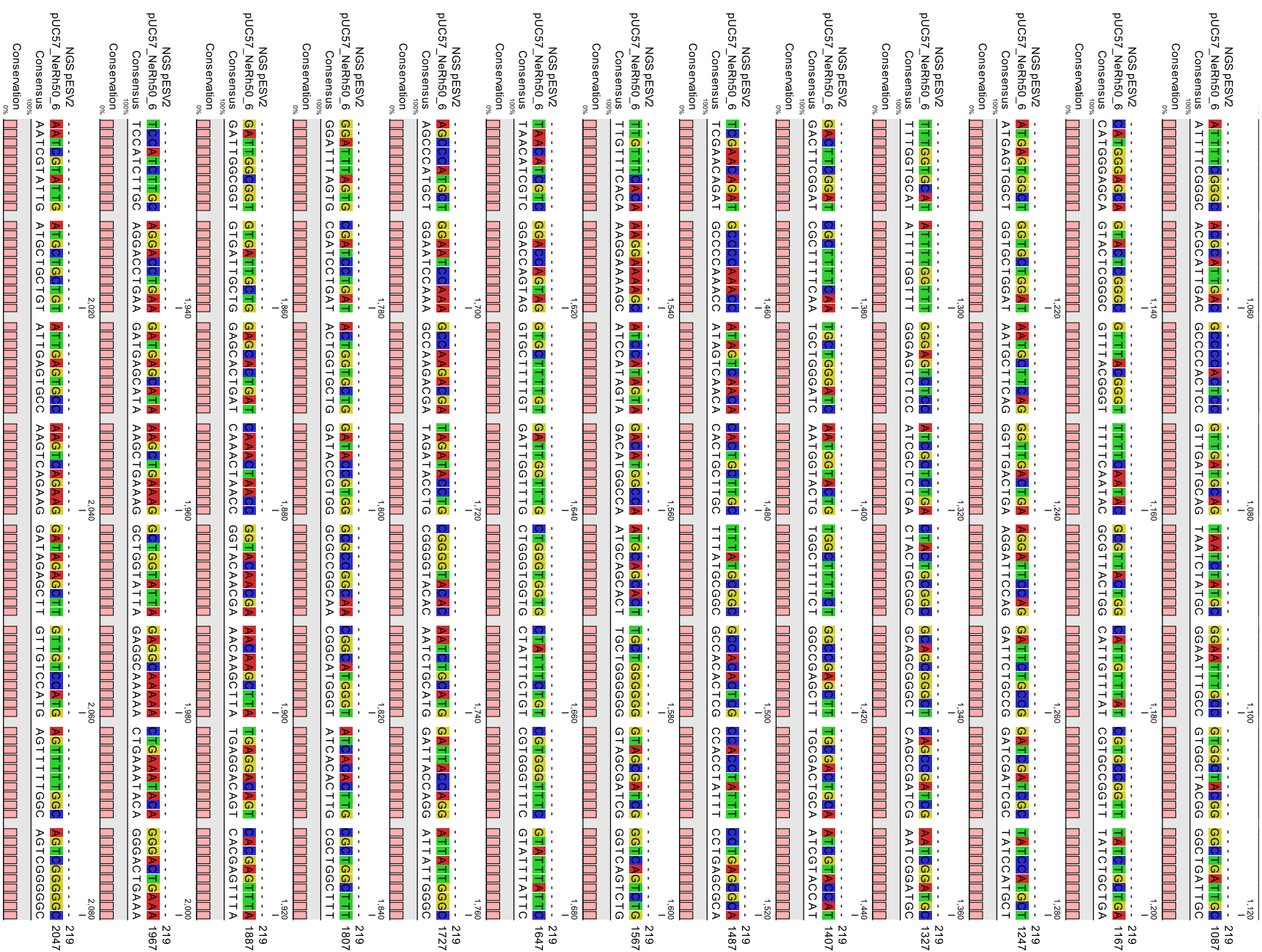
Chapter 5

Appendices

5.1 Appendix A

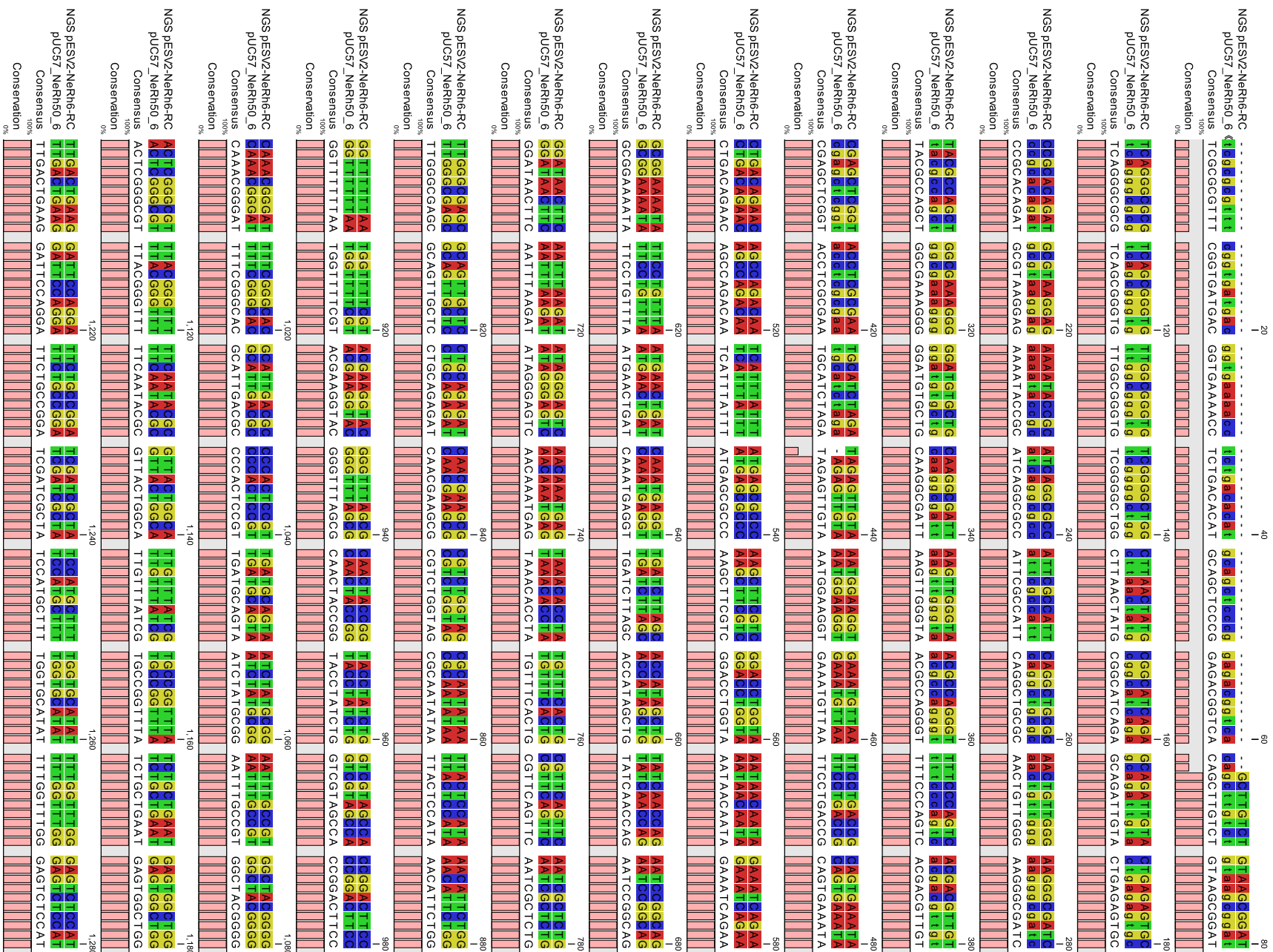
Alignment of linear nucleotides sequence of the original plasmid provided from Prof. Merrick's lab against plasmid pUC57-*NeRh50"6* from GenScript. It was aligned (100%) with the nucleotides sequence of pUC57-*NeRh50"6* provided by GenScript. Sequences were aligned by CLC Genomics Workbench 12.0 programme.

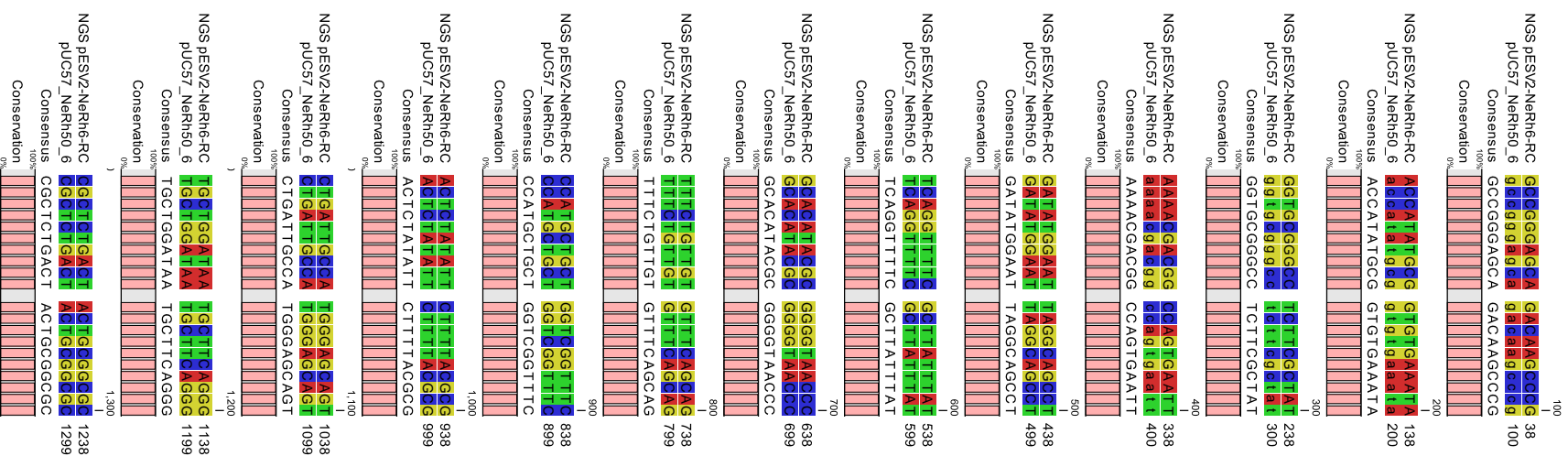


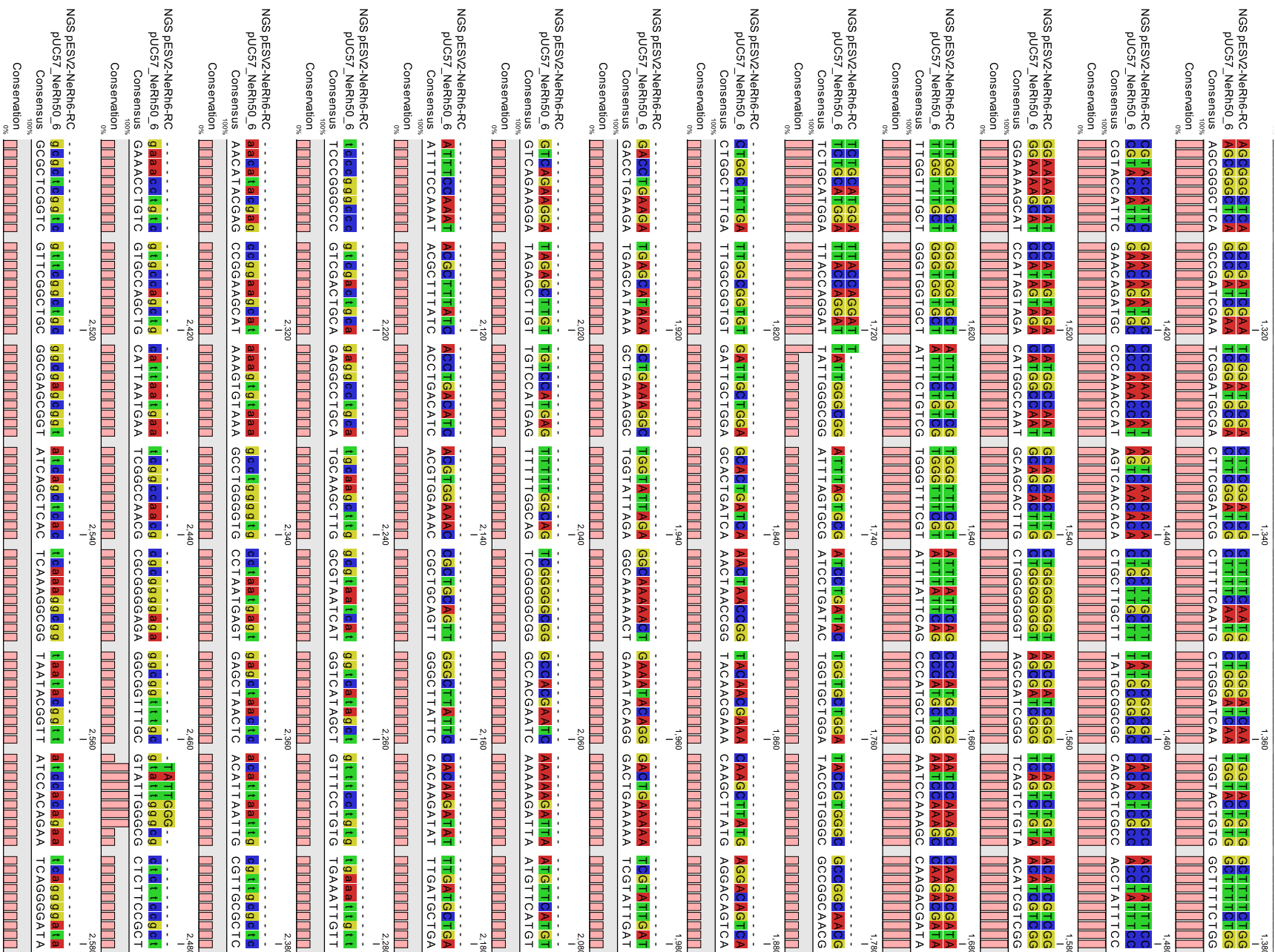


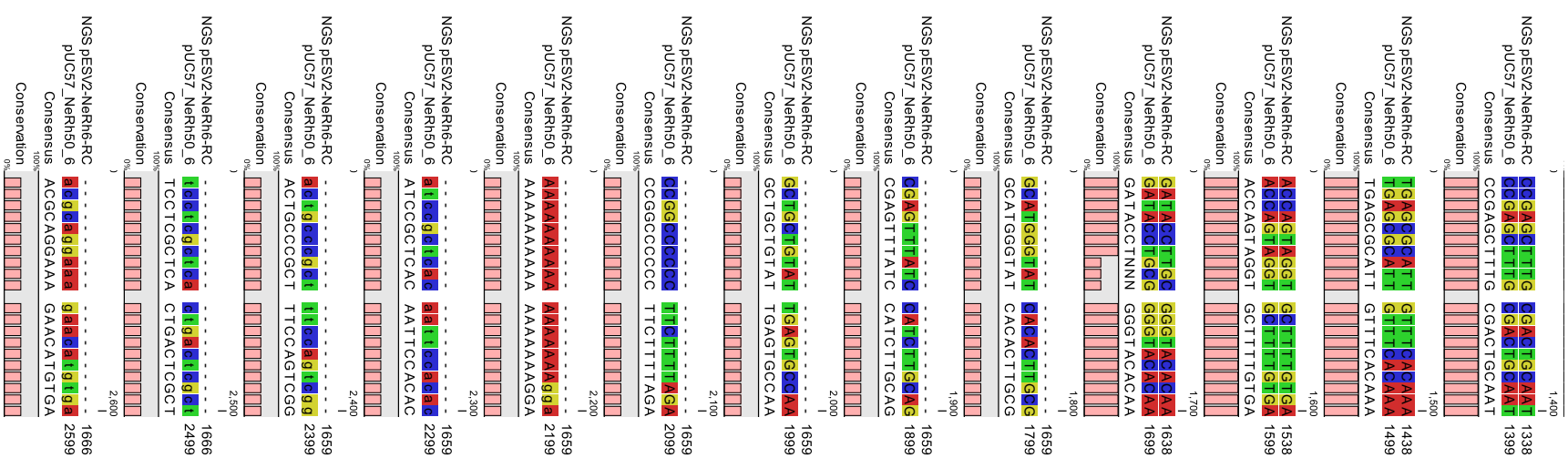
5.2 Appendix B

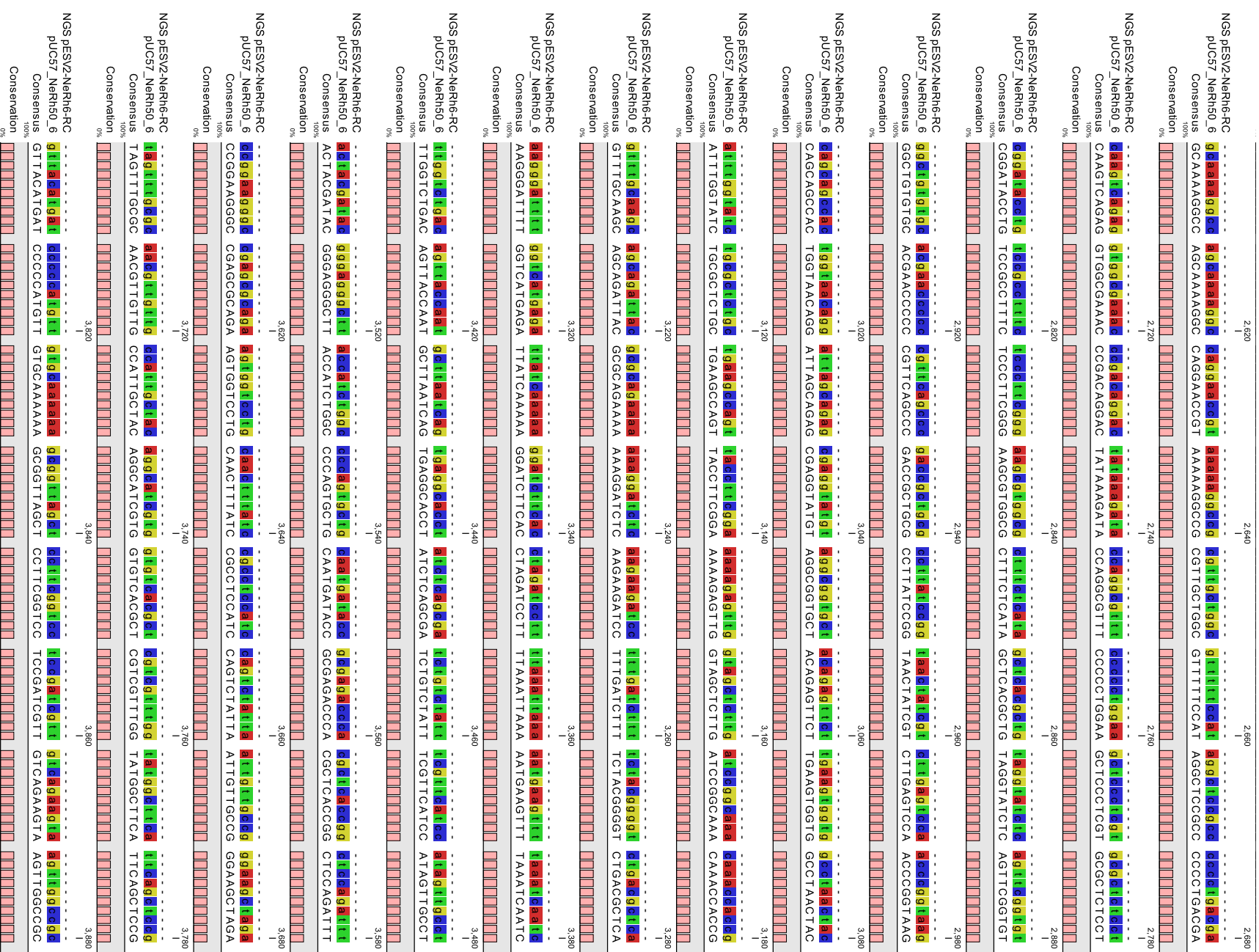
Alignment of linear nucleotides sequence of the pESV2- *NeRh50-6* against plasmid pUC57-*NeRh50"6* from GenScript. It was aligned (100%) with the nucleotides sequence of pUC57-*NeRh50"6* provided by GenScript. Sequences were aligned by CLC Genomics Workbench 12.0 programme.





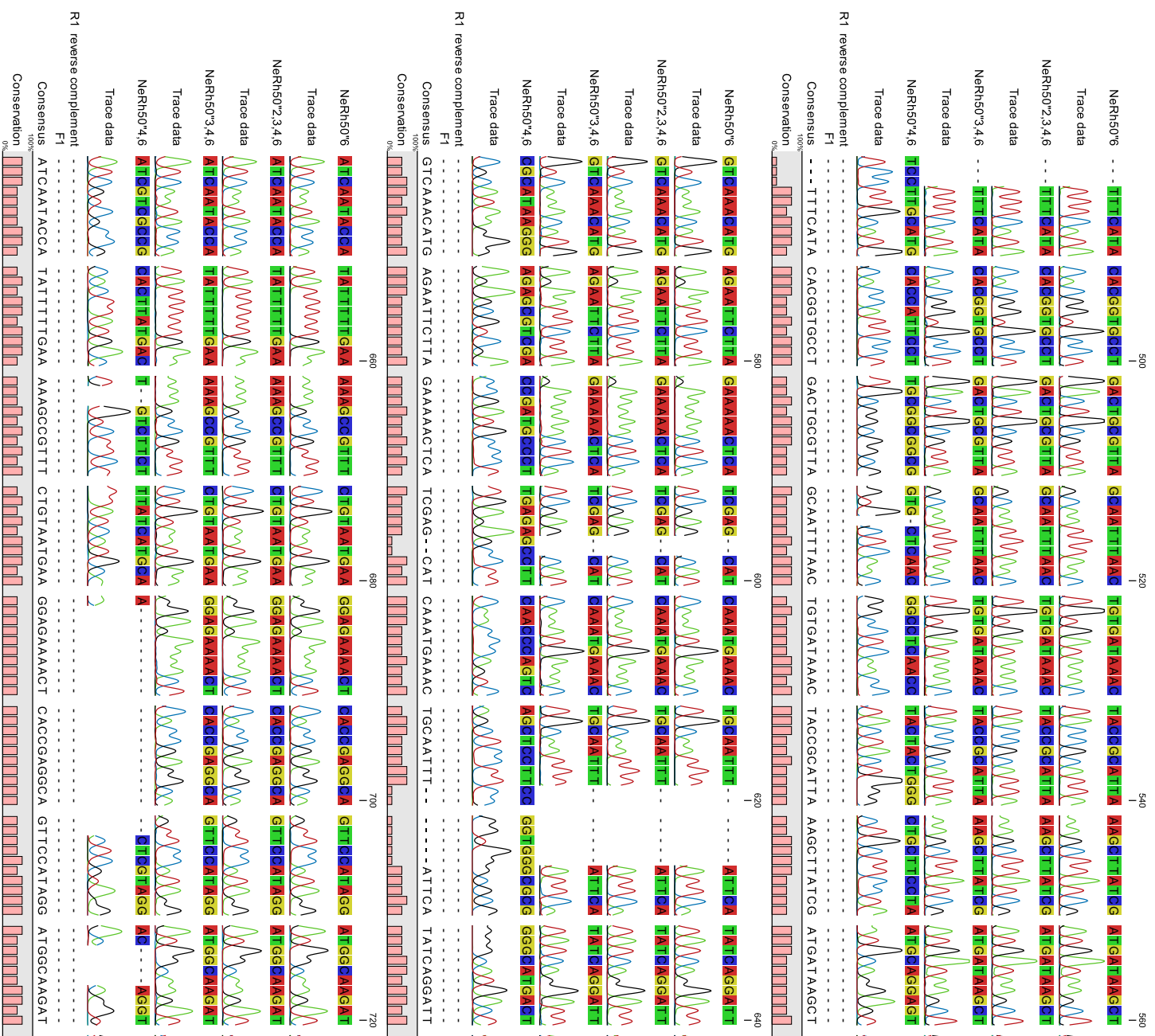


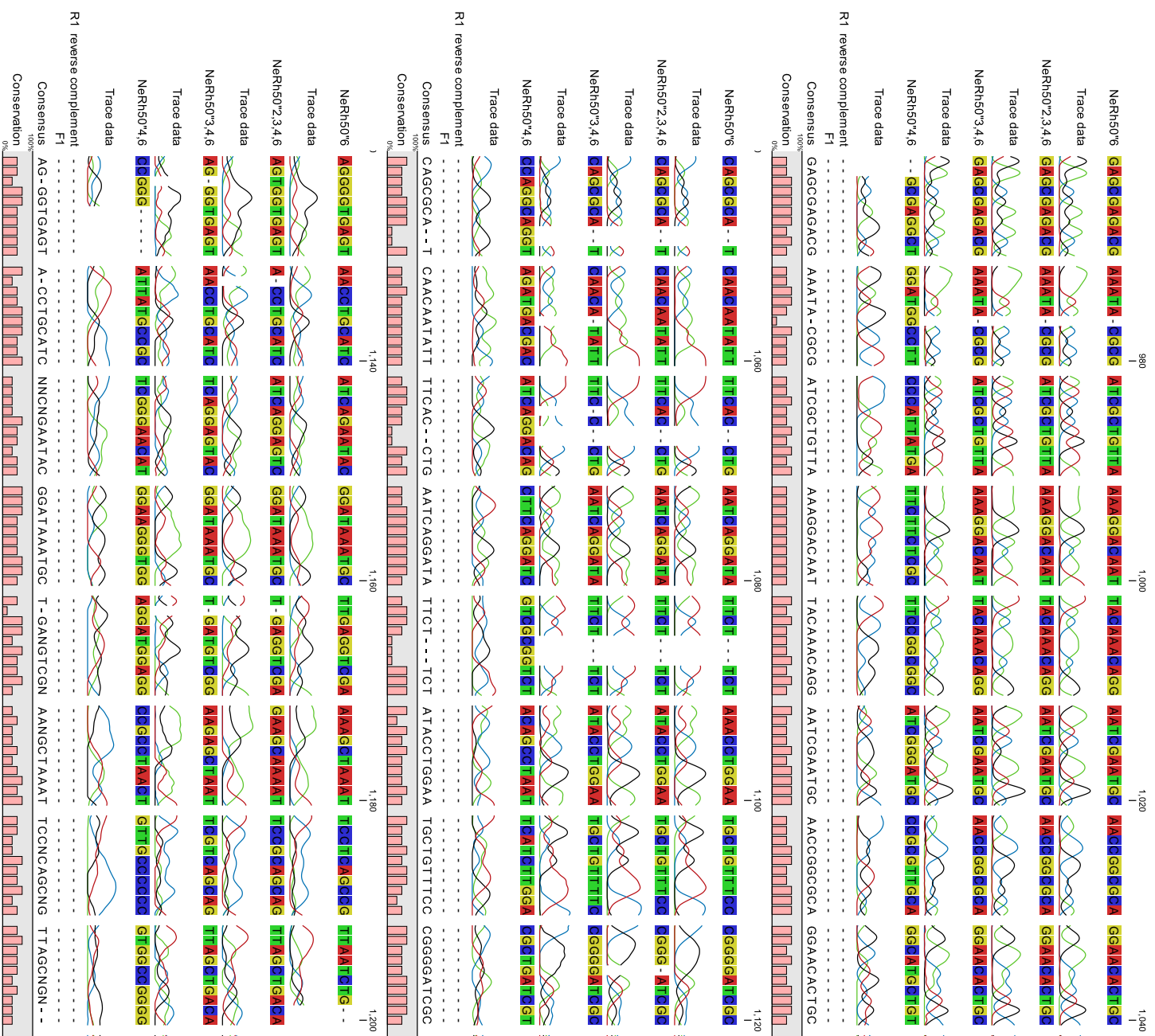


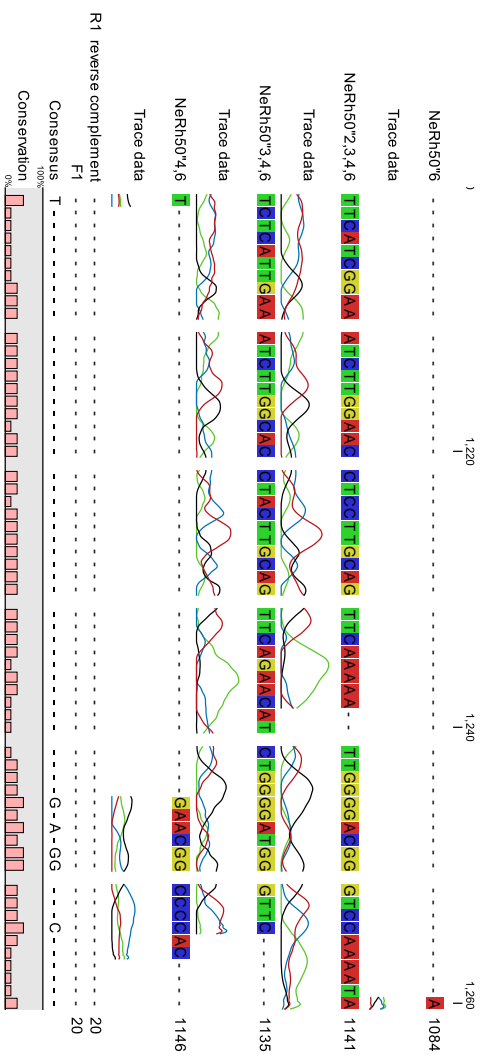


5.3 Appendix C

Sanger sequencing data with electropherogram of all constructs cDNAs: *NeRh50*"6, *NeRh50*"4, 6, *NeRh50*"3, 4, 6, and *NeRh50*"2, 3, 4, 6 were aligned by CLC Genomics Workbench 12.0 programme against: F1 forward primer/ R1 reverse primer which cover the sequence in the pET expression vector before insertion.

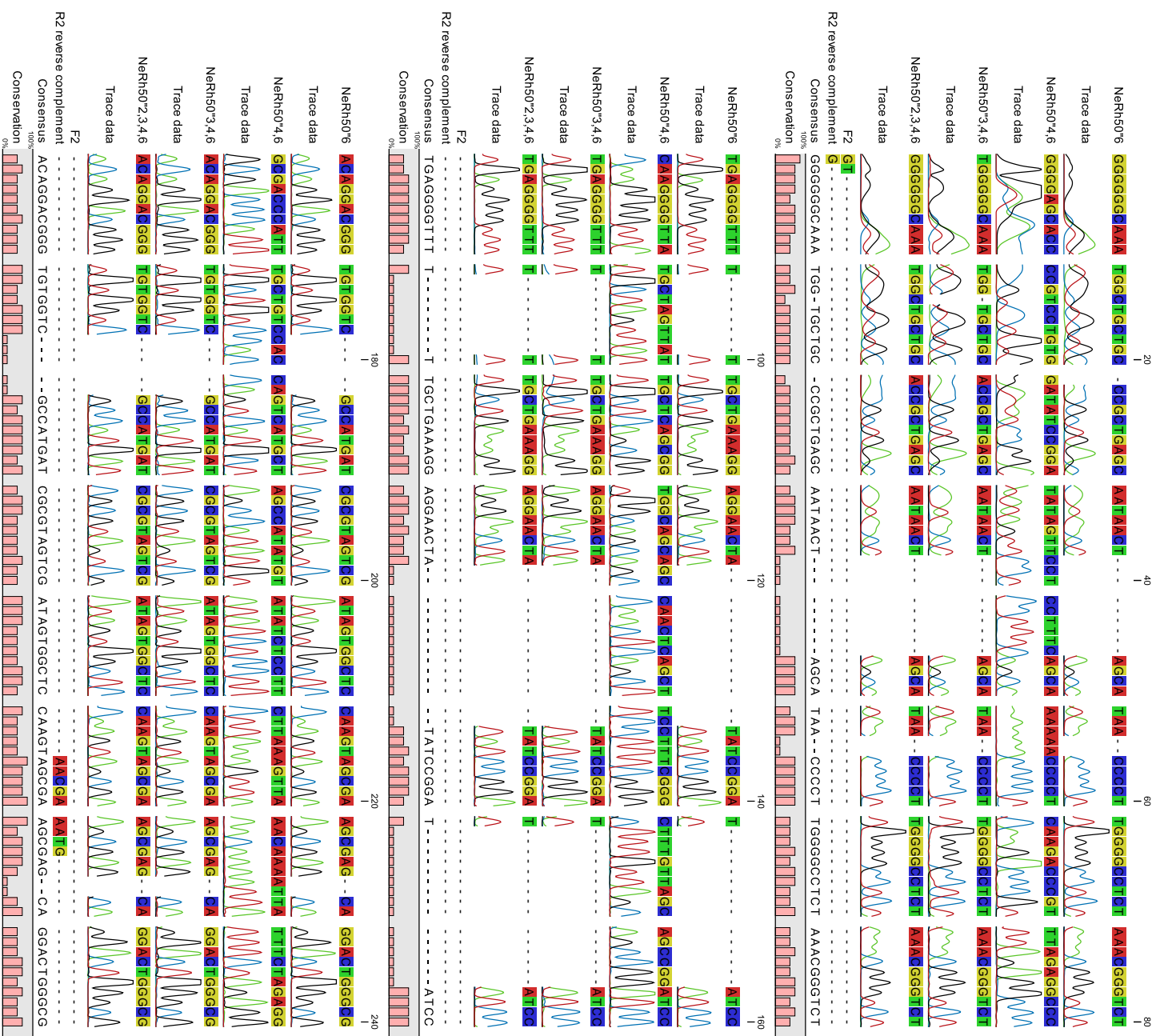


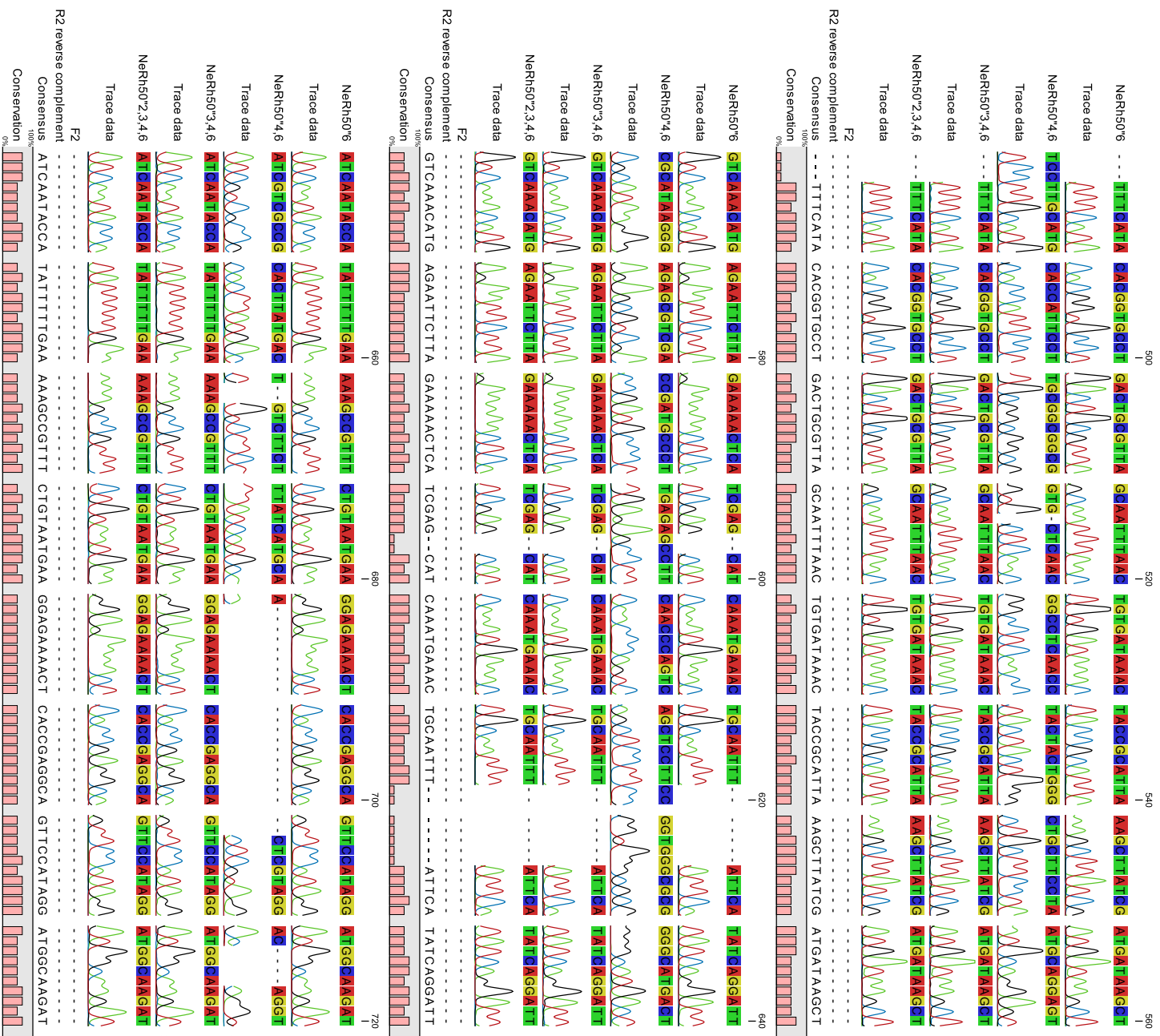


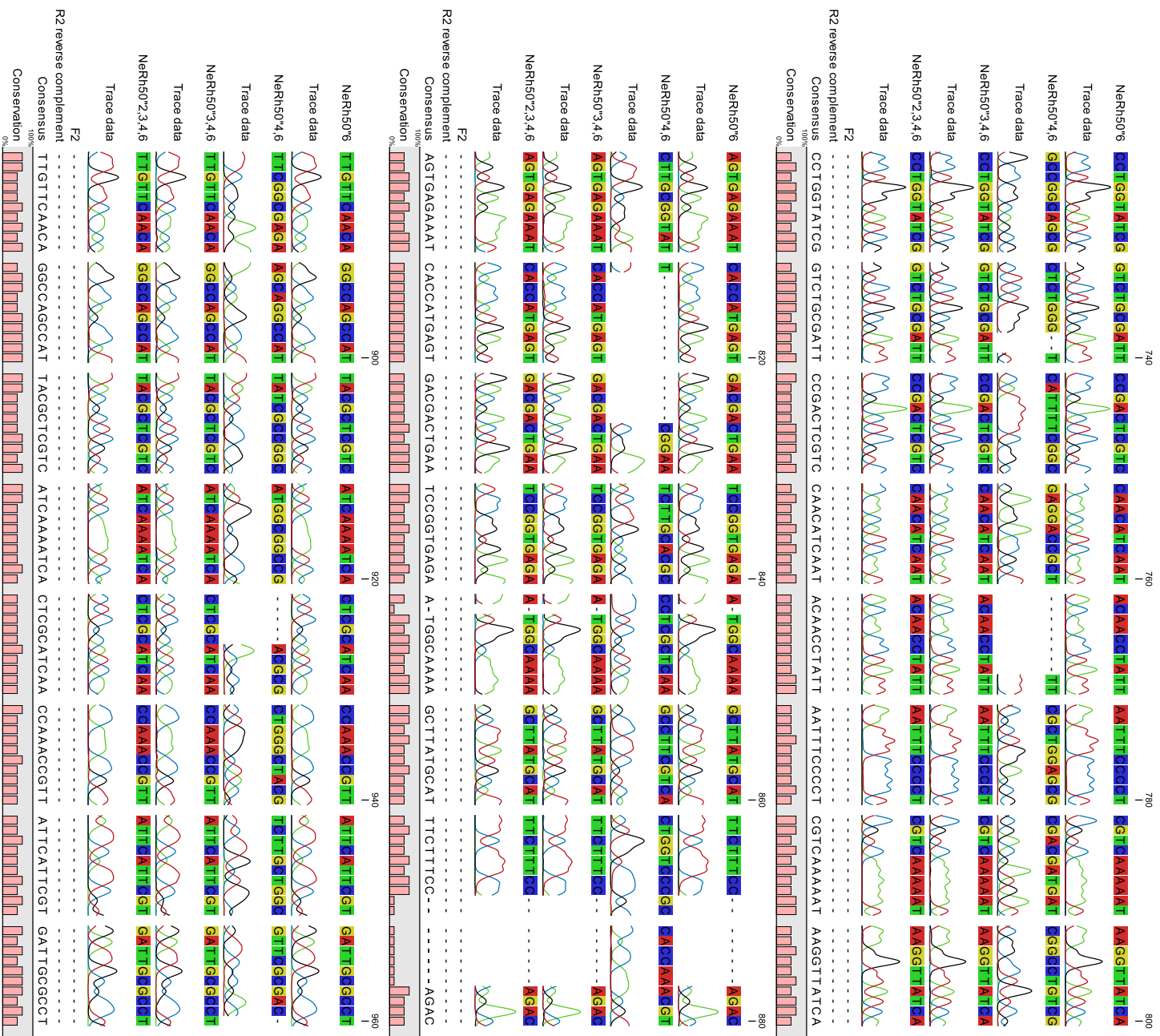


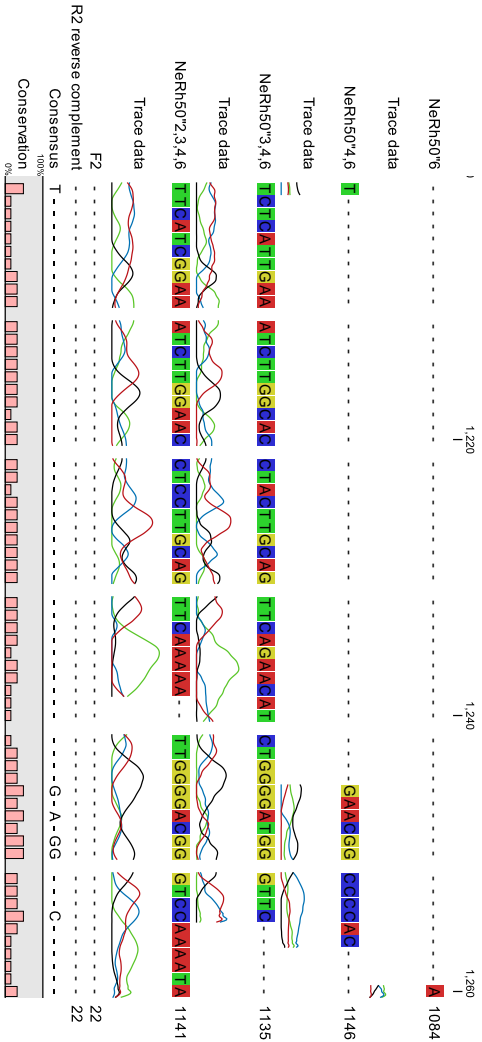
5.4 Appendix D

Sanger sequencing data with electropherogram of all constructs cDNAs: *NeRh50*"6, *NeRh50*"4, 6, *NeRh50*"3, 4, 6, and *NeRh50*"2, 3, 4, 6 were aligned by CLC Genomics Workbench 12.0 programme against: F2 forward primer/ R2 reverse primer which cover the sequence in the pET expression vector within the insertion.



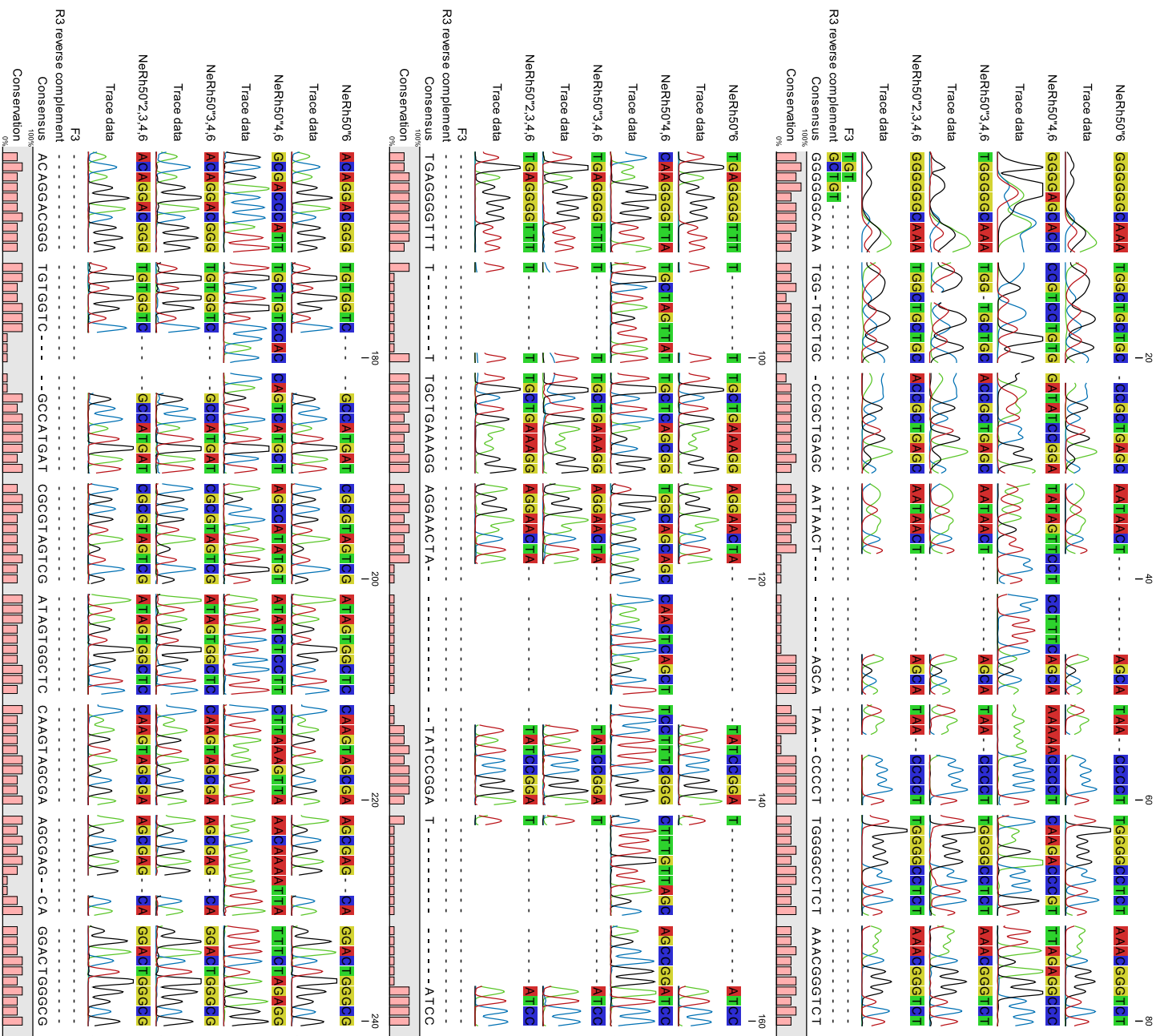


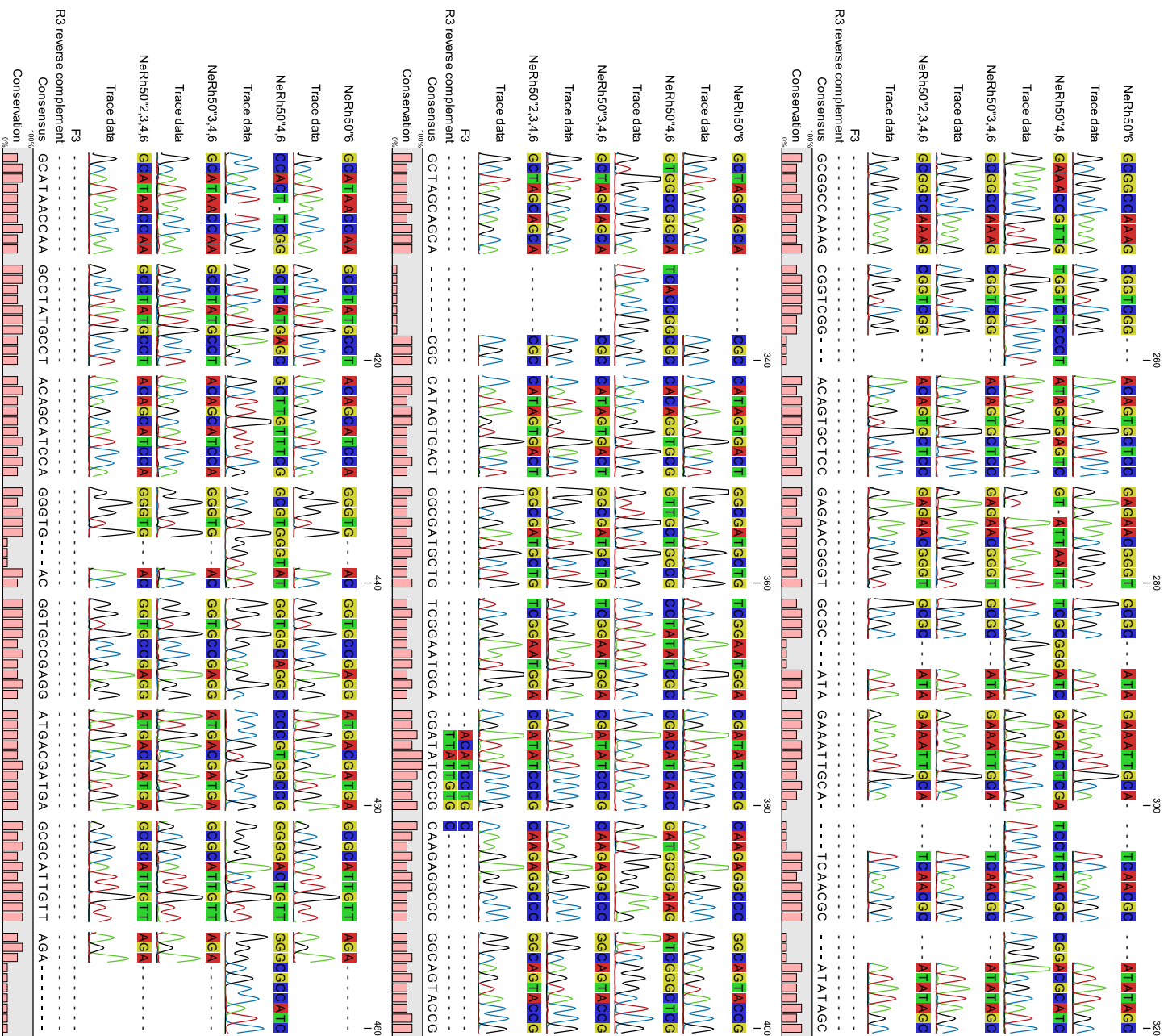


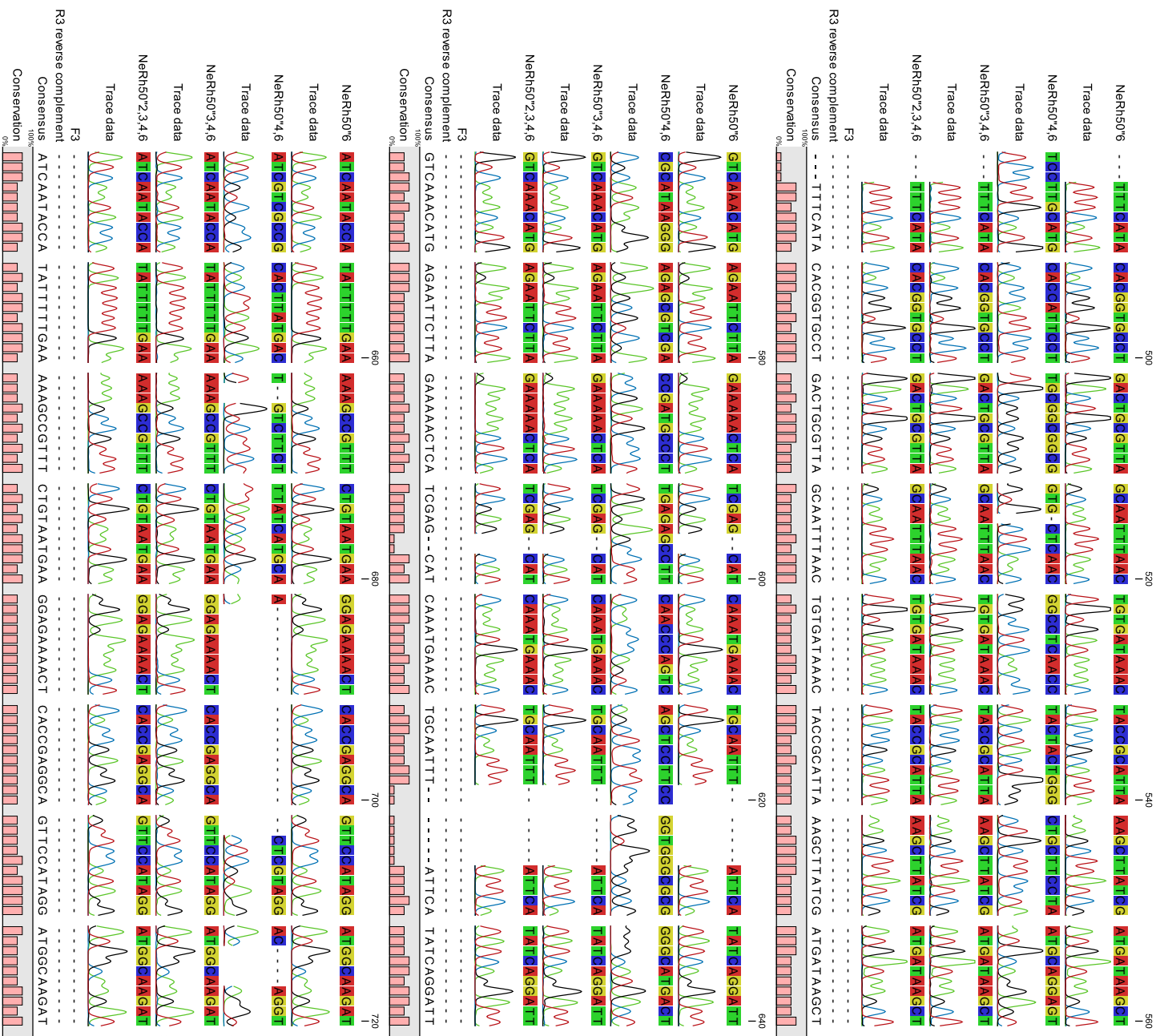


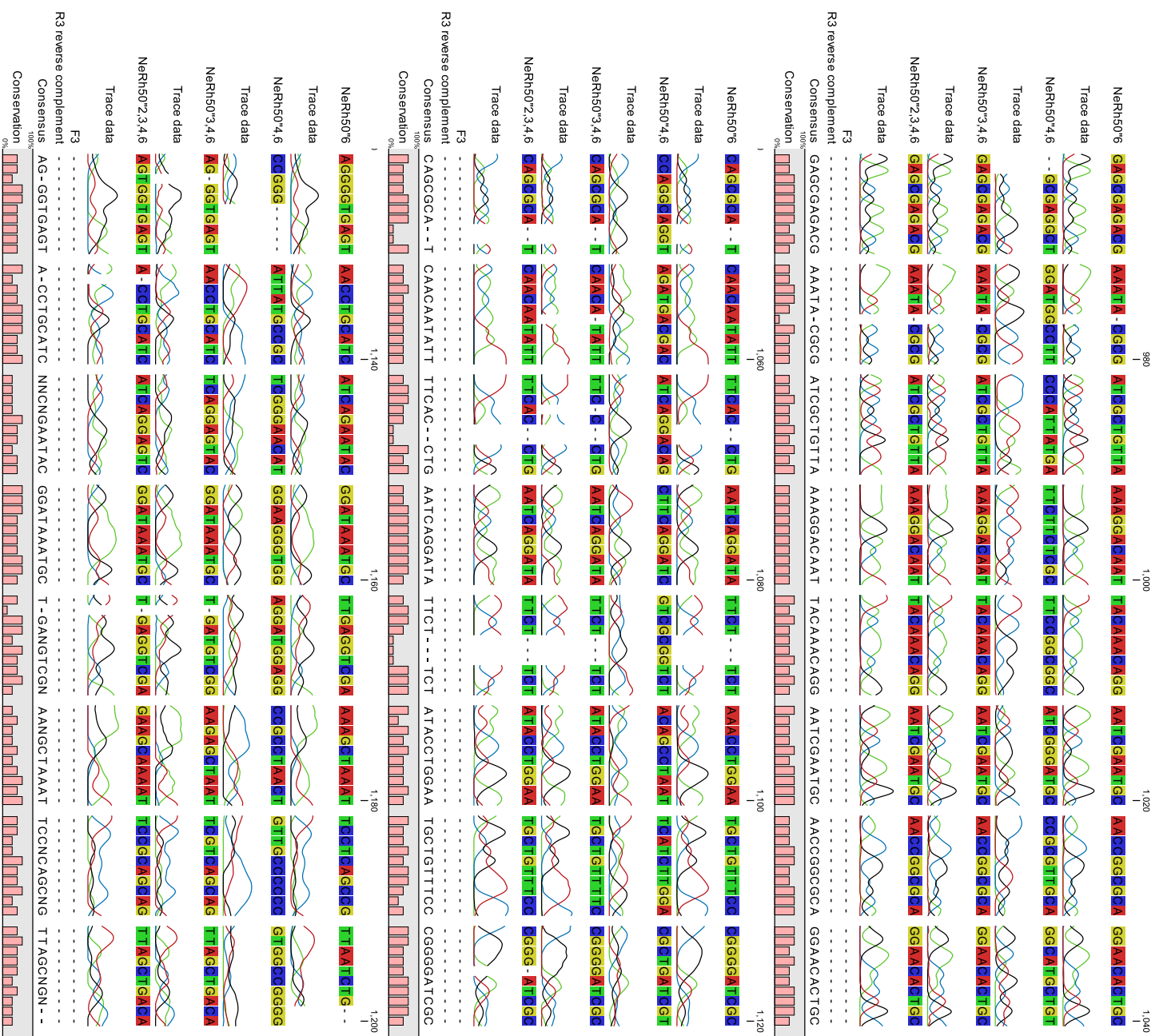
5.5 Appendix E

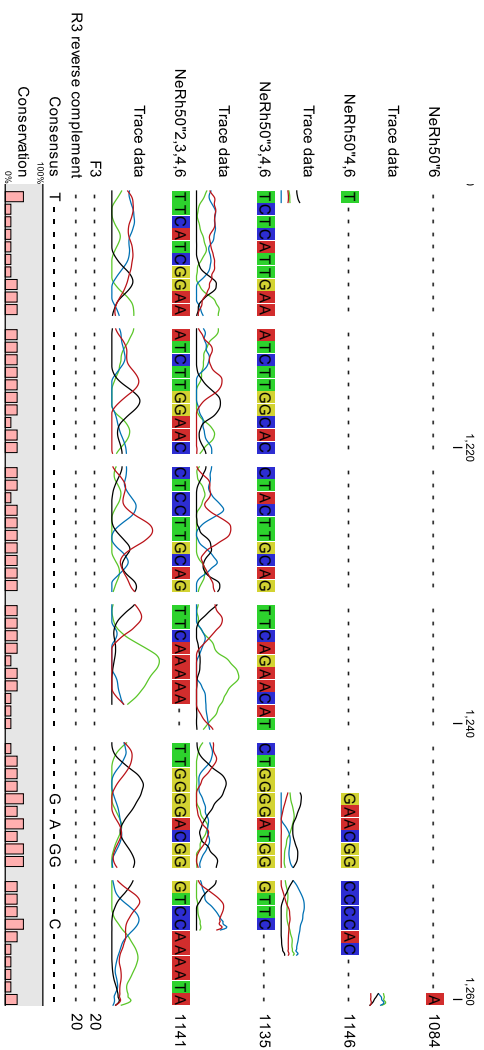
Sanger sequencing data with electropherogram of all constructs cDNAs: *NeRh50*"6, *NeRh50*"4, 6, *NeRh50*"3, 4, 6, and *NeRh50*"2, 3, 4, 6 were aligned by CLC Genomics Workbench 12.0 programme against: F3 forward primer/ R3 reverse primer which cover the sequence in the pET expression vector within the insertion.





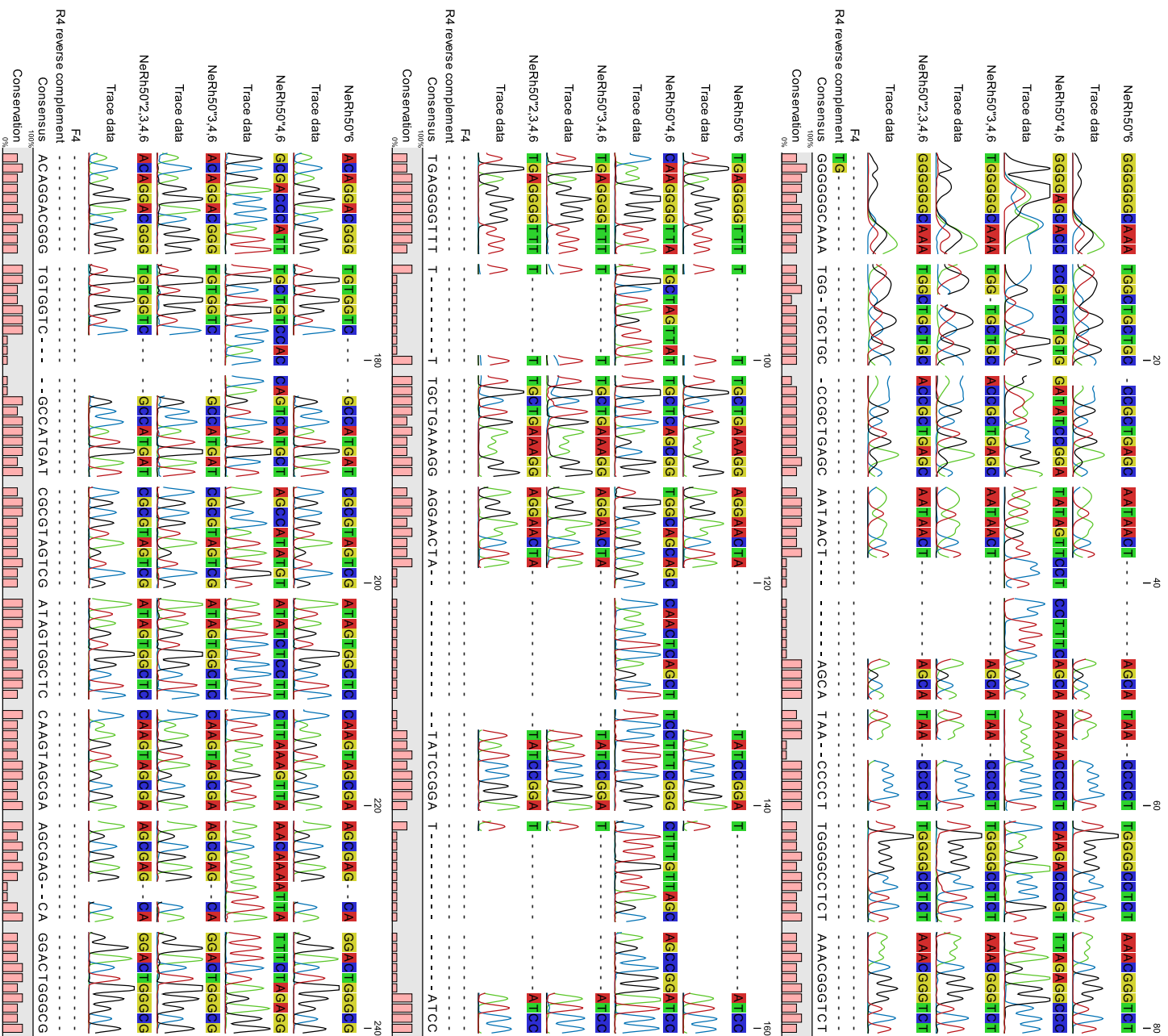


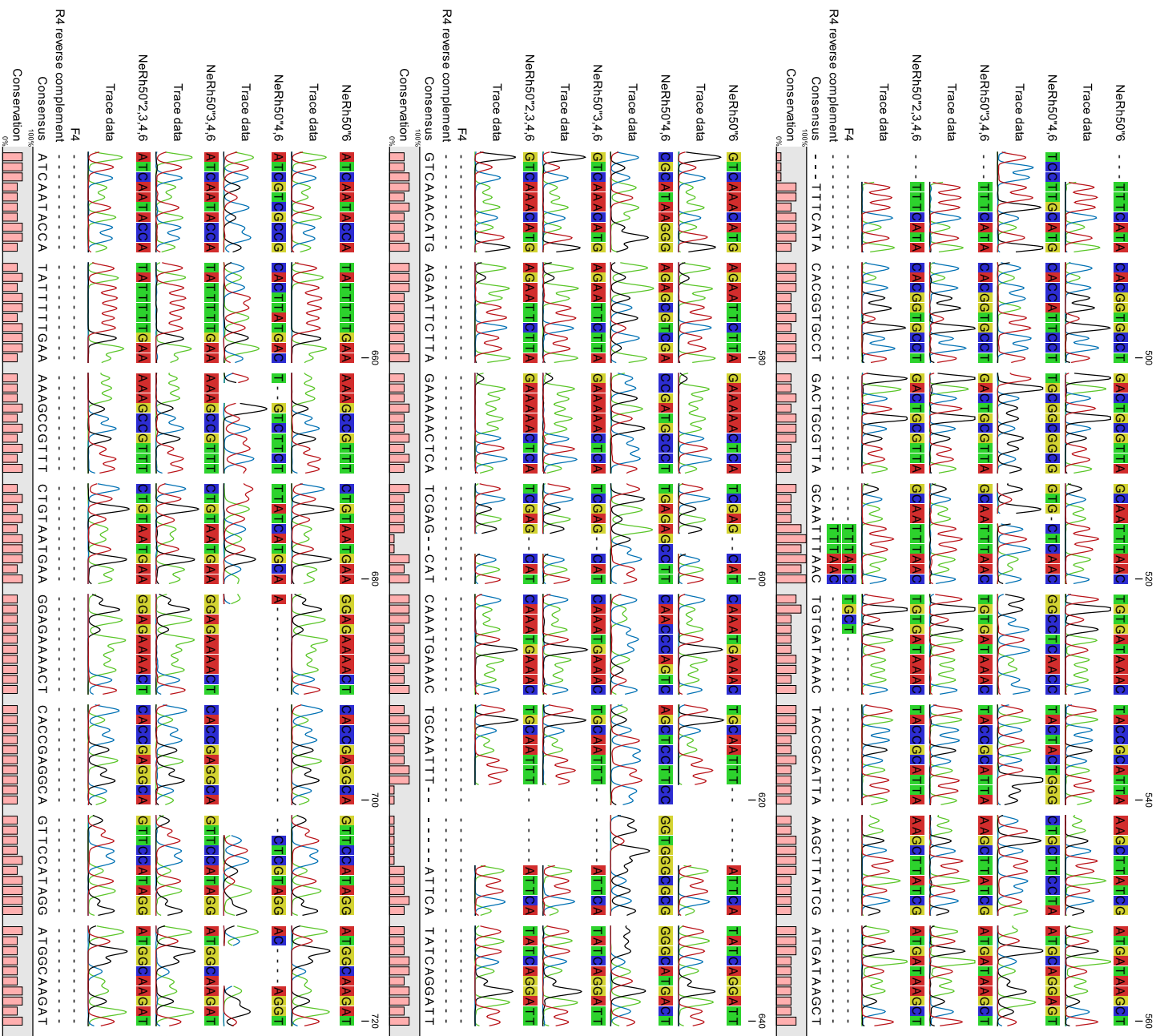


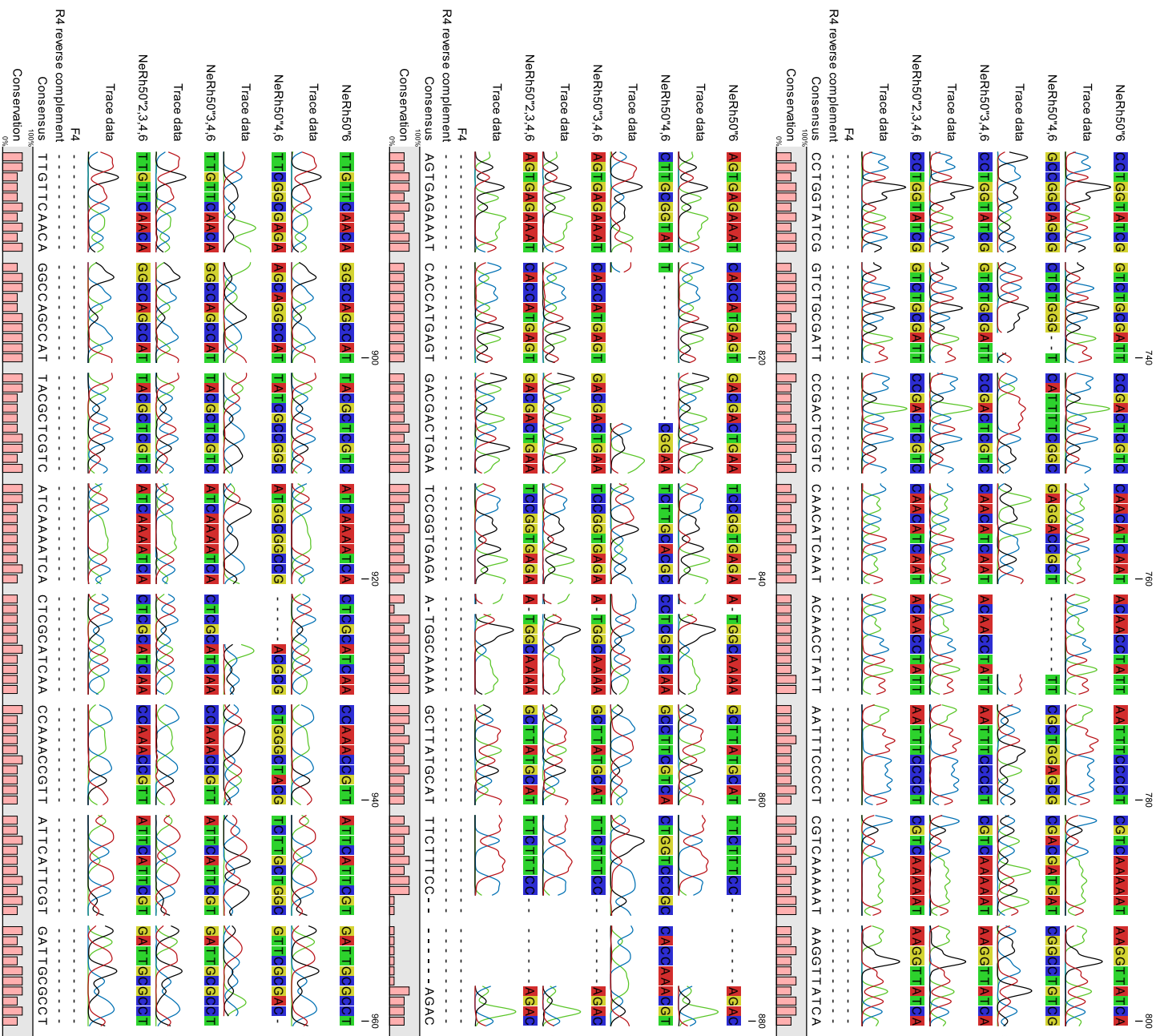


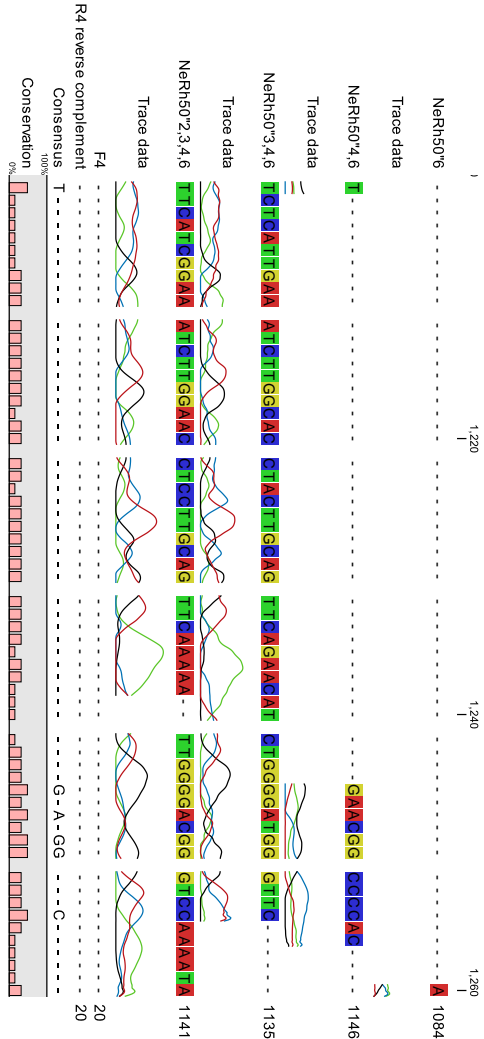
5.6 Appendix F

Sanger sequencing data with electropherogram of all constructs cDNAs: *NeRh50*"6, *NeRh50*"4, 6, *NeRh50*"3, 4, 6, and *NeRh50*"2, 3, 4, 6 were aligned by CLC Genomics Workbench 12.0 programme against: F4 forward primer/ R4 reverse primer which cover the sequence in the pET expression vector within the insertion.



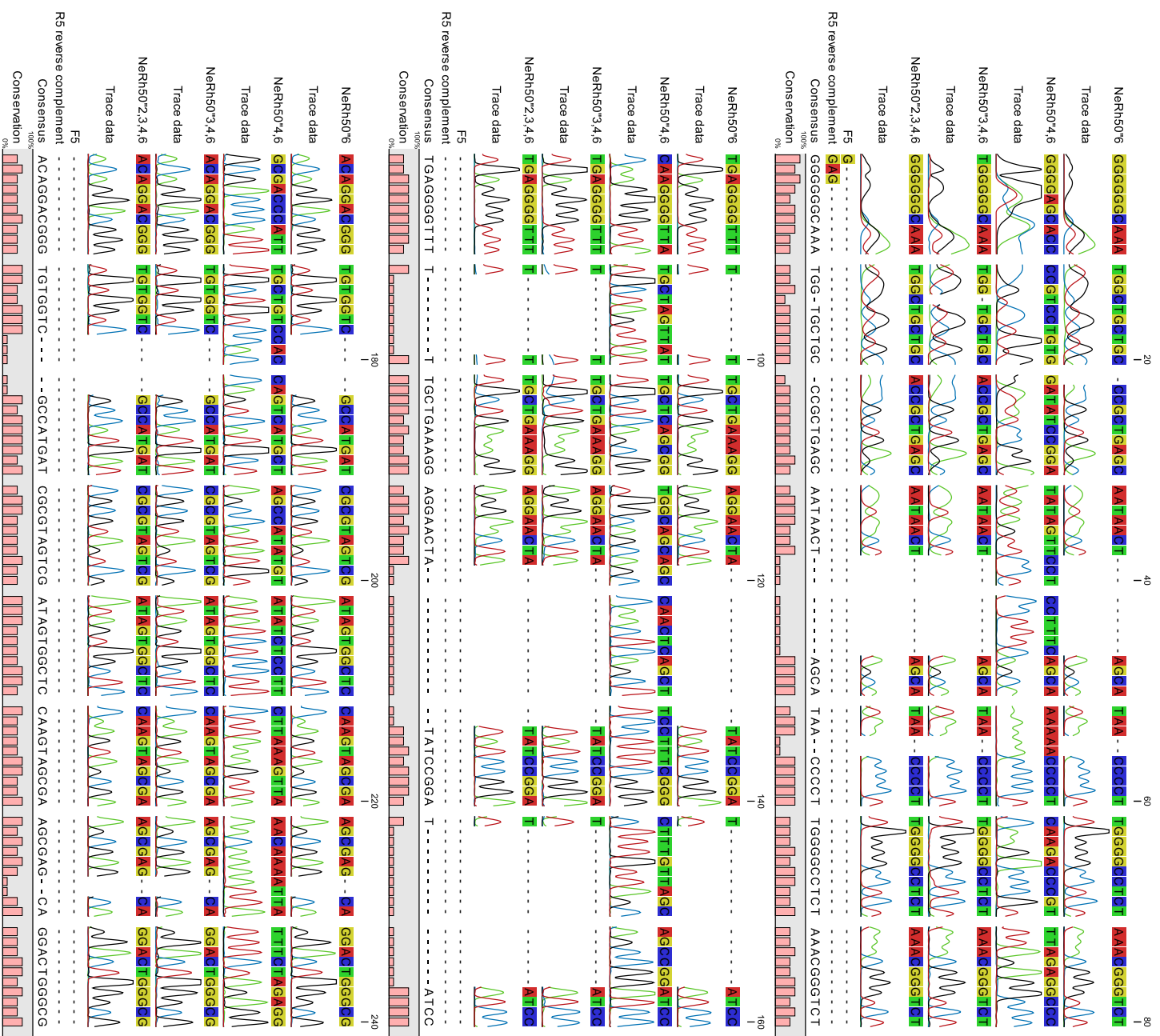


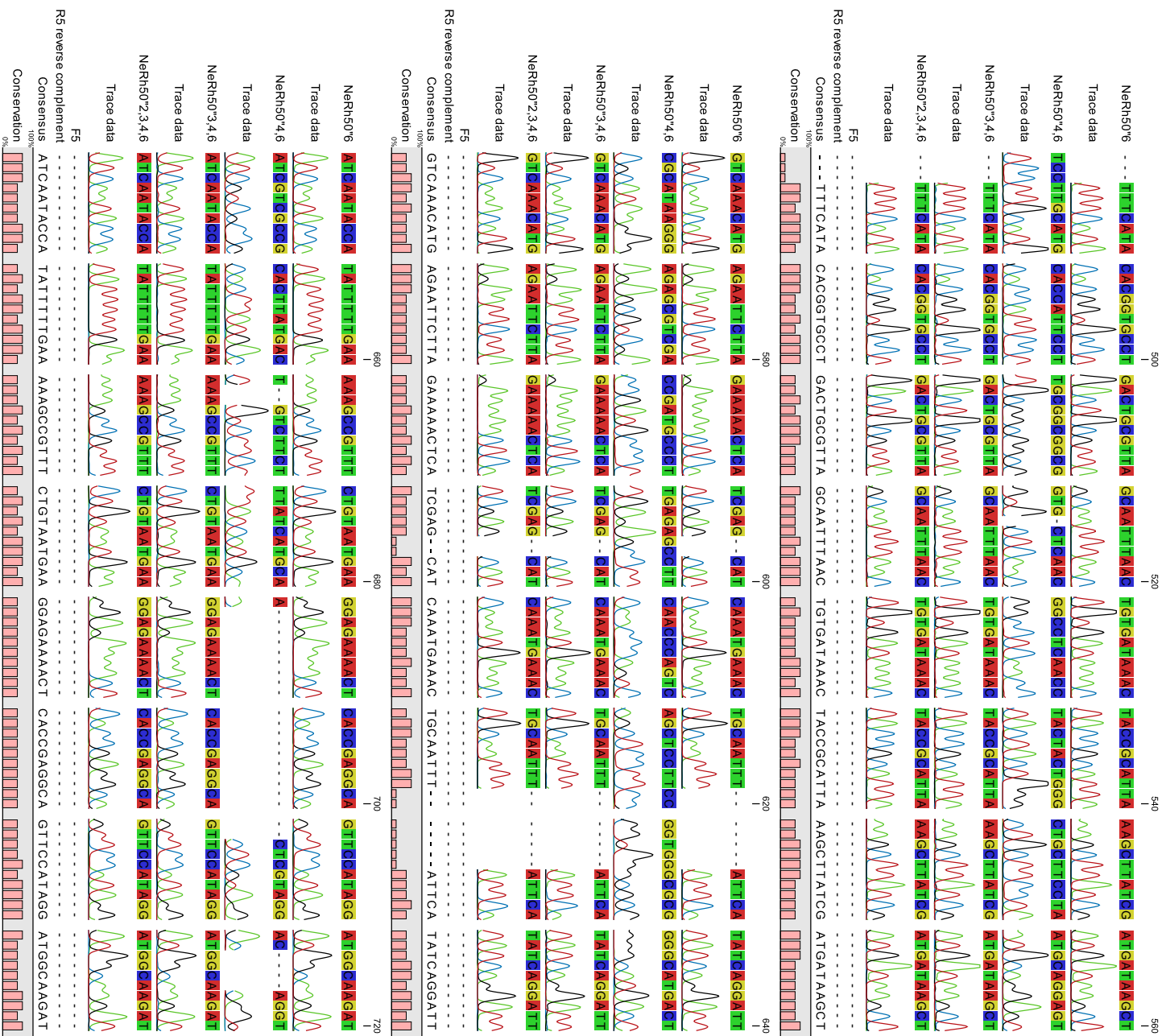


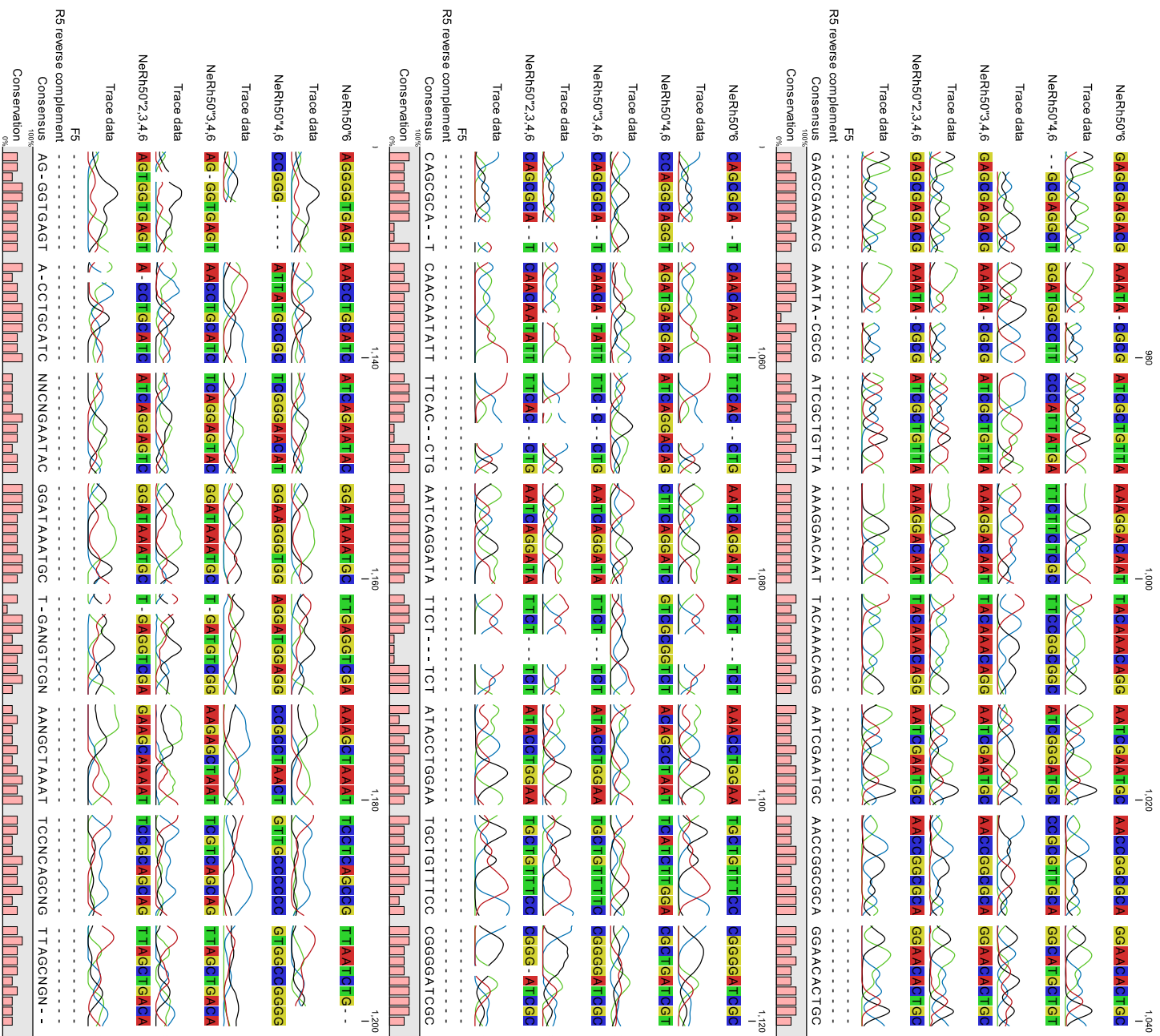


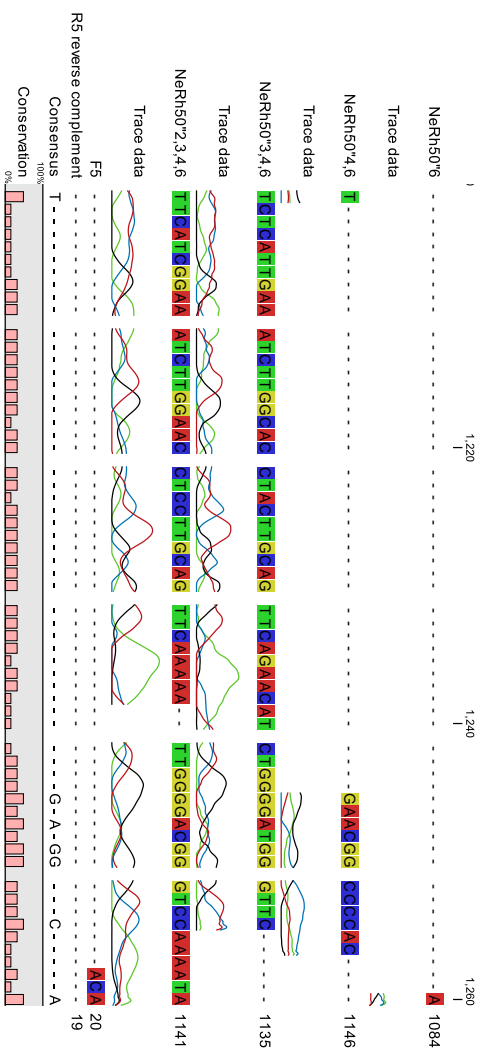
5.7 Appendix G

Sanger sequencing data with electropherogram of all constructs cDNAs: *NeRh50*"6, *NeRh50*"4, 6, *NeRh50*"3, 4, 6, and *NeRh50*"2, 3, 4, 6 were aligned by CLC Genomics Workbench 12.0 programme against: F5 forward primer/ R5 reverse primer which cover the sequence in the pET expression vector within the insertion.



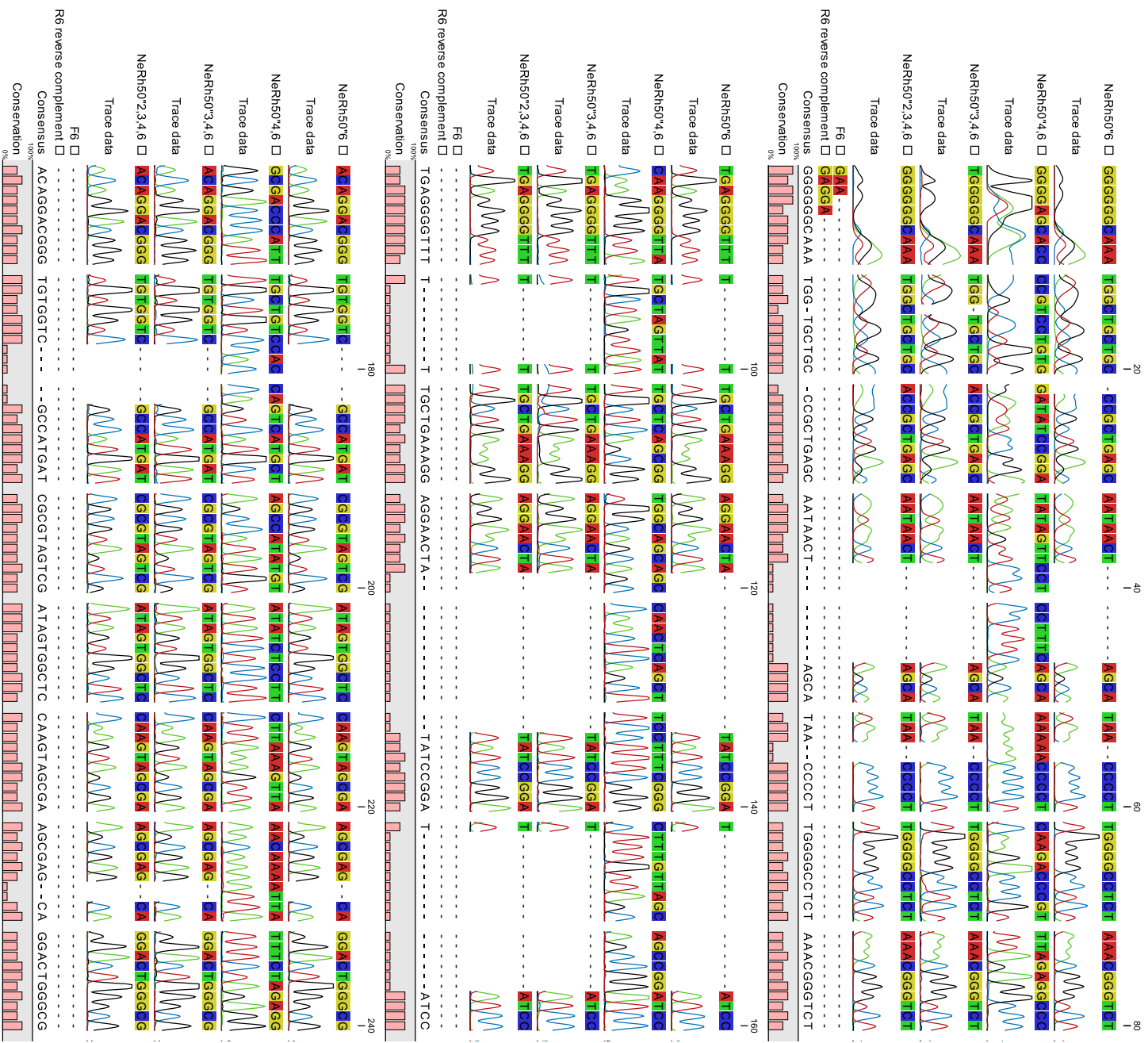


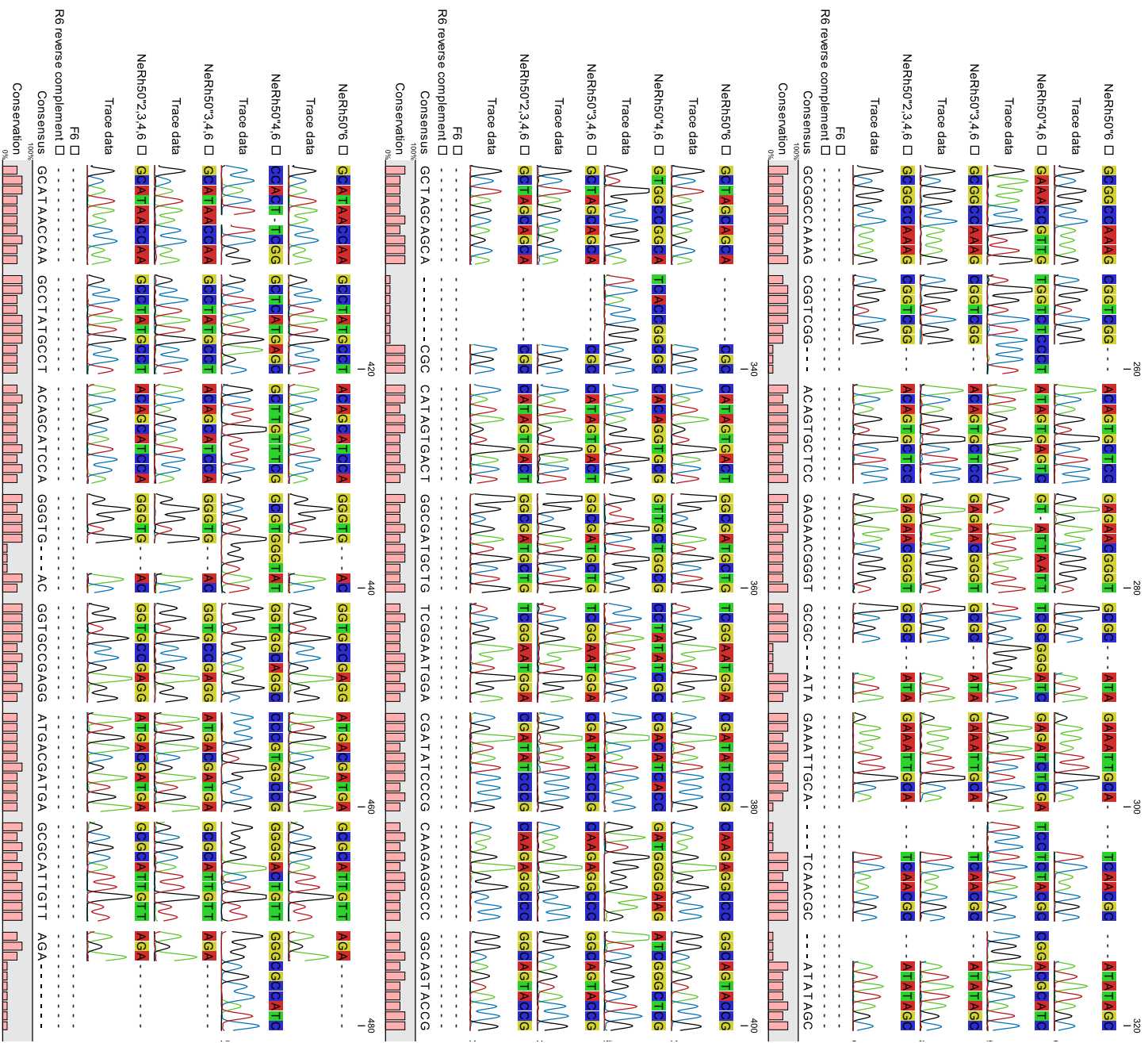


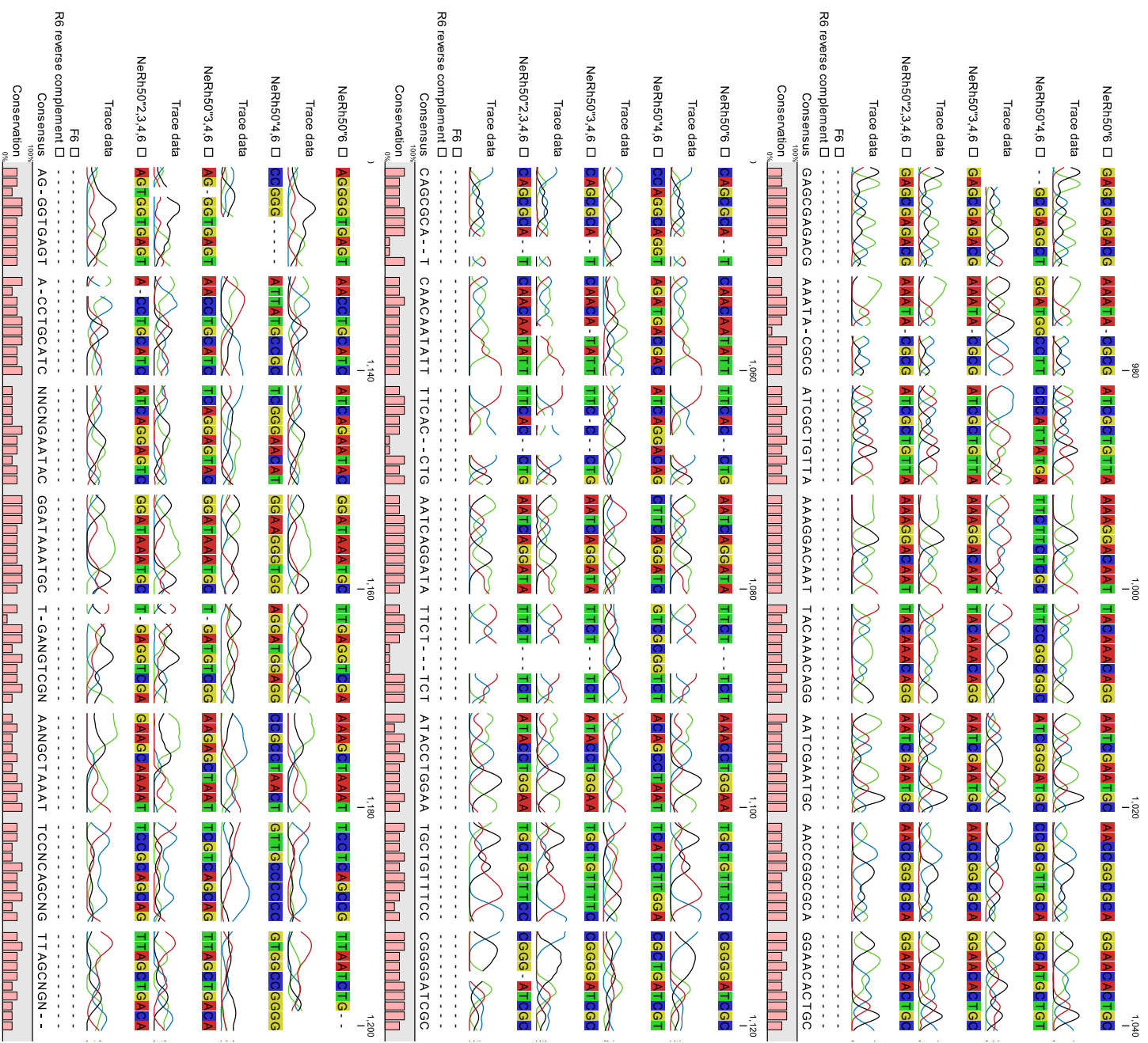


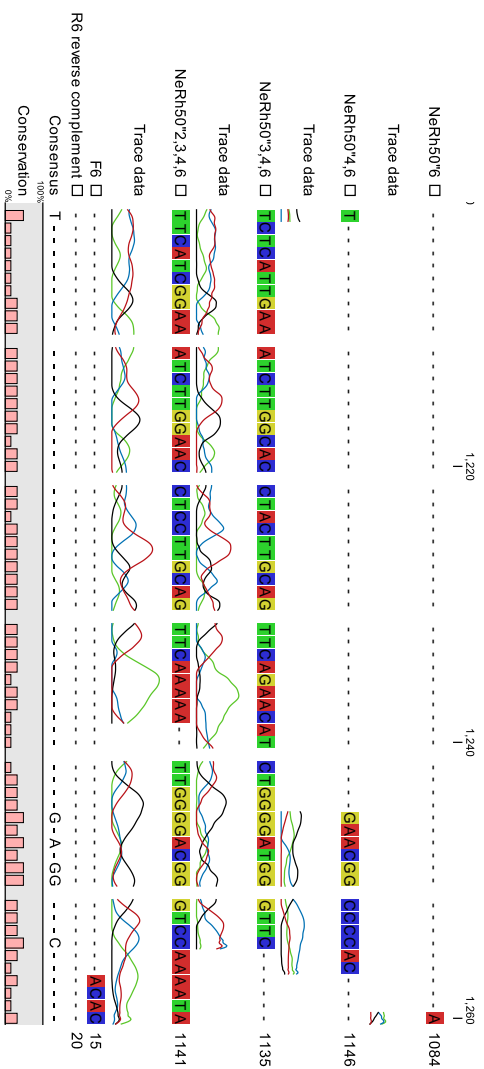
5.8 Appendix H

Sanger sequencing data with electropherogram of all constructs cDNAs: *NeRh50*"6, *NeRh50*"4, 6, *NeRh50*"3, 4, 6, and *NeRh50*"2, 3, 4, 6 were aligned by CLC Genomics Workbench 12.0 programme against: F6 forward primer/ R6 reverse primer which cover the sequence in the pET expression vector within the insertion.









5.9 Appendix I

Sanger sequencing data with electropherogram of all constructs cDNAs: *NeRh50*"6, *NeRh50*"4, 6, *NeRh50*"3, 4, 6, and *NeRh50*"2, 3, 4, 6 were aligned by CLC Genomics Workbench 12.0 programme against: F7 forward primer/ R7 reverse primer which cover the sequence in the pET expression vector after the insertion.

